Daniela Ferreira Garcia Rodrigues

Volatile Organic Compounds for the Detection of Bladder Cancer: an *In Vitro* Metabolomic Approach

Dissertation thesis for the Master Degree in
Analytical Clinical and Forensic Toxicology

Dissertação do 2º Ciclo de Estudos conducente ao grau de
Mestre em Toxicologia Analítica Clínica e Forense

Trabalho realizado sob orientação de Professora Doutora Márcia Cláudia Dias de Carvalho e co-orientação de Doutora Maria Paula do Amaral Alegria Guedes de Pinho

October 2016

This work received financial support from the European Union (FEDER funds POCI/01/0145/FEDER/007728) and National Funds (FCT/MEC, Fundação para a Ciência e Tecnologia and Ministério da Educação e Ciência) under the Partnership Agreement PT2020 UID/MULTI/04378/2013.

This study is a result of the project NORTE-01-0145-FEDER-000024, supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement (DESignBIOtecHealth - New Technologies for three Health Challenges of Modern Societies: Diabetes, Drug Abuse and Kidney Diseases), through the European Regional Development Fund (ERDF).

ACKNOWLEDGEMENTS

This thesis could not be accomplished without the help, support and guidance of some people that accompanied me during this year and that contribute not only for the improvement of my work skills but also to my personal growth.

Firstly, I would like to express special thanks to Prof. Márcia Carvalho and Dr. Paula Guedes (my supervisors) for all their help and guidance during this year. I am also grateful for their valuable advices, their patience and, specially, for believing in me and in my work. Without their support, this thesis wouldn't be possible. I'm also grateful to Prof. Maria de Lourdes Bastos (my co-supervisor) for the support, the knowledge shared and the encouragement words given throughout the Master course.

I extend my thanks to Prof. Carmen Jerónimo, from the Instituto Português de Oncologia do Porto (IPO), who kindly provided the Bladder cancer cell lines and for her medical expertise in the research and in the published papers. I also thank to Prof. Rui Henriques and Prof. Luís Belo for their contribution in the review papers.

To Prof. António Barros (University of Aveiro) who kindly provided the statistical software Monte Carlo cross validation (MCCV).

I'm grateful to Ana Margarida for her friendship, patience and always being available for sharing her experience, knowledge, and also when I needed help or looked for advice during this work. I extend my gratitude to Joana Pinto for her support, advice and, so much needed, help in the processing of the data and during all multivariate statistical analysis. Her assistance and knowledge were of major importance for the completion of this thesis.

To Márcia Monteiro for spending some of her time helping me, advising and checking if everything was OK.

To Ana Sofia Oliveira who always offered her help and assistance when I needed, I appreciate everything. To all the rest of the people and friends from the lab, Rita, Cátia, Margarida, Patrícia, Maria Enea, Rosário and Jorge, thank you for the help, partnership and good company.

To my special friend Ni, whose friendship has grown the most during this last months, thank you for your help, company, support, especially in those hard-working days, for the laughs, the discussion of ideas, the good lunches and, of course, the delicious cakes and pancakes that you never forgot to bring and share. To my other two special ladies from the Master course, Maria and Armanda, with whom I also shared great moments, laughs, knowledge and last but not least, wonderful sushi lunches.

To my “biochemist” friends, who I bring with me since the first day at University, thank you for your important friendship, the support and the good memories we have created over the years. I’ll never forget the dinners, celebrations, the philosophical discussions whether about work or life.

I’m also very grateful to my “older” best friends Ana, Daniela and Inês because no matter how different our paths were, we always managed to find time to reunite and share all the successes, failures, worries and joys. Thank you for the constant support, all the “good lucks” and the “everything is ok?”. They sure helped me during the years and I hope they continue to.

To you, my special M, I am grateful for the partnership, the inspiration, the care and for always being there, listening and supporting. Thank you for believing in me and being my constant reminder that I will succeed.

Finally, to my family, thank you so much for the unconditional love, trust, patience, for always believing in me and in my success, and for giving me the freedom of following my dreams and making my own choices. Thank you for the encouragement words, the valuable advices, the efforts and sacrifices you’ve made in order to assure that nothing was missing.

To all of you, a sincere Thank You.

PUBLICATIONS

Articles in international peer-reviewed journals

Rodrigues, D., Jerónimo, C., Henrique, R., Belo, L., de Lourdes Bastos, M., de Pinho, P. G. and Carvalho, M. (2016), Biomarkers in bladder cancer: A metabolomic approach using *in vitro* and *ex vivo* model systems. **Int. J. Cancer**, 139: 256–268. doi:10.1002/ijc.30016

Rodrigues, D., Monteiro, M., Jerónimo, C., Henrique, R., Belo, L., de Lourdes Bastos, M., de Pinho, P. G. and Carvalho, M. (2016), Renal Cell Carcinoma: A critical analysis of metabolomic biomarkers emerging from current model systems. **Transl Res.** doi: 10.1016/j.trsl.2016.07.018

ABSTRACT

Metabolomics is a burgeoning field that involves the detection of a great variety of compounds which are present in several matrices, including cells, tissues or biofluids. As metabolites are end products of the biochemical pathways, alterations in their levels reflect alterations in the normal function of the organism. Therefore, metabolomics has been implemented in several clinical areas, mainly because a particular disease or pathological condition may present a unique metabolic profile. This is of major importance for disease diagnosis, prognosis and monitoring. Additionally, metabolite levels can be measured through analysis of urine or blood, whose collection is less invasive than biopsies. Despite its endless applications, metabolomics has shown to be particularly useful in the search of diagnosis cancer biomarkers. Its application in cancer research relies on the fact that cancer cells suffer a metabolic reprogramming to balance their energetic needs to support the rapid tumour growth. Indeed, there are several metabolomic studies reporting alterations in metabolites levels in many types of cancers, namely bladder cancer (BC). BC is one of the most common and fatal malignancies of the urinary system, whose incidence continues to rise. Current diagnosis methods are extremely invasive and poor sensitive. Thus, they often detect BC at advanced stages where prognosis is dismal and the probability of treatment success and patients' survival rate are low. For these reasons, it is evident the necessity in identifying new biomarkers and to develop earlier diagnosis strategies.

The work presented in this thesis provides a proof-of-concept for the *in vitro* prediction of BC from volatile organic compounds (VOCs) analysis of the extracellular medium of bladder cells (both cancer and normal), using Headspace-Solid Phase Microextraction/Gas Chromatography-Mass Spectrometry (HS-SPME/GC-MS) technique. Aiming at the investigation of the potential of volatiles as new biomarkers for BC detection, as well as for the classification of the disease according to histological grade and subtype, this study comprised the analysis of three bladder cancer cell lines (5637, J82 and Scaber) and one normal cell line (SV-HUC-1). Two different pHs (pH 2 and pH 7) were also evaluated in order to determine the best pH for VOCs extraction to be used in future *in vitro* metabolomic studies. After statistical analysis, it was observed that VOCs collected from the bladder cells exometabolome, at both pHs, enable the separation of BC cell lines from the normal one, as well as the discrimination between low-grade and high-grade BC and between two different BC histological subtypes. These results are encouraging as they demonstrate the biomarker potential of volatiles for BC diagnosis. Nevertheless, the metabolic pathways that lead to VOCs production are not yet deeply elucidated, hindering

a complete comprehensive interpretation of the results. Therefore, validation studies and application of the method to clinical samples (particularly urine) from BC patients are required not only to improve knowledge about VOCs metabolism, but also to evaluate and confirm the translatability of these VOCs to clinical practice.

Keywords: Bladder cancer (BC); Metabolomics; Volatile Organic Compounds (VOCs); Biomarkers; HS-SPME/GC-MS

RESUMO

A metabolómica é uma área da ciência que tem crescido exponencialmente, e que envolve a deteção de uma grande variedade de compostos presentes em várias matrizes, incluindo células, tecidos ou fluídos biológicos. Sendo os metabolitos produtos finais das vias metabólicas, alterações nos seus níveis refletem alterações no normal funcionamento do organismo. Assim, a metabolómica tem sido implementada em várias áreas da biomedicina, principalmente porque uma determinada doença ou condição patológica podem apresentar um perfil metabólico específico. Isto é bastante importante na medida em que pode ajudar no prognóstico, diagnóstico e monitorização de doenças. Além disso, os níveis dos metabolitos podem ser avaliados através da análise de urina ou sangue, cuja recolha é menos invasiva do que as biópsias. Apesar das suas infinitas aplicações, a metabolómica tem demostrado ser útil na procura de biomarcadores de diagnóstico de cancro. A sua aplicação na oncologia assenta no facto das células cancerígenas sofrerem uma reprogramação metabólica de forma a equilibrarem as suas necessidades energéticas, para, assim, manterem o rápido crescimento do tumor. De facto, já existem bastantes estudos metabolómicos que revelam alterações nos níveis de metabolitos em vários tipos de cancro, nomeadamente o cancro da bexiga. Este cancro é uma das neoplasias mais comuns e fatais do trato urinário, cuja incidência continua a aumentar. Os métodos de diagnóstico atuais são extremamente invasivos e pouco sensíveis. Por isso, o cancro da bexiga é, na maioria das vezes, diagnosticado em estádios já avançados, cujo prognóstico é péssimo, sendo baixos a probabilidade de um tratamento ter sucesso e o tempo de sobrevivência dos doentes. Por estas razões, há uma necessidade evidente em identificar novos biomarcadores e em desenvolver métodos que permitam o diagnóstico mais precoce do cancro em questão.

O trabalho apresentado nesta tese permite obter uma “prova de conceito” para o diagnóstico *in vitro* do cancro da bexiga, através da análise de compostos orgânicos voláteis (COVs) do meio extracelular de células da bexiga (tanto cancerígenas como normais), usando a técnica da microextração em fase sólida no *headspace* (HS-SPME) seguida de cromatografia gasosa acoplada à espetrometria de massa (GC-MS). Com o objetivo de investigar potenciais voláteis como novos biomarcadores de diagnóstico do cancro da bexiga, bem como a classificação deste cancro de acordo com o grau e subtipo, este estudo incluiu a análise de três linhas de bexiga cancerígenas (5637, J82 e Scaber) e uma linha celular normal (SV-HUC-1). Foram avaliados dois pHs diferentes (pH 2 e pH 7) para determinar qual deles seria o melhor para a extração de voláteis, de forma a ser aplicado em estudos *in vitro* metabolómicos futuros. Depois de ser feita uma análise

estatística dos resultados, observou-se que os voláteis extraídos do exometaboloma das células da bexiga, em ambos os pHs, permitiram a separação das linhas cancerígenas da linha normal, bem como a discriminação entre cancro de baixo e alto grau e entre dois subtipos diferentes de cancro da bexiga. Estes resultados são promissores na medida em que mostram o potencial dos voláteis como biomarcadores de diagnóstico deste cancro. Contudo, as vias metabólicas que levam à produção de COVs ainda não estão totalmente compreendidas, impedindo uma interpretação mais precisa dos resultados. Portanto, são precisos mais estudos para a validação dos COVs e aplicação deste método em amostras biológicas (particularmente urina) de doentes com cancro da bexiga para não só melhorar o nosso conhecimento sobre o metabolismo dos COVs, mas também para avaliar e confirmar a translação dos mesmos para a clínica.

Palavras-chave: Cancro da bexiga; metabolómica; compostos orgânicos voláteis (COVs); biomarcadores; HS-SPME/GC-MS

TABLE OF CONTENTS

INDEX OF FIGURES

INDEX OF TABLES

LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
BC	Bladder cancer
BSA	Bovine Serum Albumin
CDF	Computable Document Format
CESs	Carboxylesterases
CR	Classification rate
CRC	Colorectal cancer
Cu^{+}	Cuprous ion
$CuSO_4 5H_2O$	Copper (II) sulphate pentahydrate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
EDTA	Ethylenediaminetetraacetic acid
EI	Electron ionization
ES	Effect size
ES_{SE}	Effect size standard error
FBS	Fetal Bovine Serum
GC	Gas Chromatography
GSH	Glutathione
HCl	Hydrochloric acid
HG	High-grade
HMBD	Human Metabolome Database
HPLC	High Performance Liquid Chromatography
HS	Headspace
HS-SPME	Headspace-Solid Phase Microextraction
IPO	Instituto Português de Oncologia
IS	Internal standard
IUPAC	International Union of Pure and Applied Chemistry

KEGG	Kyoto Encyclopaedia of Genes and Genomes
$KNaC_4H_4O_6$	Potassium sodium tartrate
LC	Liquid Chromatography
LG	Low-grade
LV	Latent variable
M	Molar
m/z	Mass-to-charge ratio
MCCV	Monte Carlo cross validation
MEM	Minimum Essential Medium
min.	Minutes
MS	Mass spectrometry
Na_2CO_3	Sodium carbonate
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NIST	National Institute of Standards and Technology
NMR	Nuclear Magnetic Resonance
OPLS-DA	Orthogonal partial least squares-discriminate analysis
Par	Pareto
PC1	First principal component
PC2	Second principal component
PCA	Principal component analysis
PDMS/DVB	Polydimethylsiloxane/Divinylbenzene
PLS-DA	Partial least squares-discriminate analysis
ppm	Parts per million
PTFE	Polytetrafluoroethylene
Q^2	Goodness of prediction or prediction power
QC	Quality control
R^2	Variance explained
RI	Retention index
ROC	Receiver operating characteristic
RSD	Relative standard deviation
RT	Retention time
SCC	Squamous cell carcinoma
sens.	Sensitivity
SMR	Standard reference materials
spec.	Specificity

SPME	Solid Phase Microextraction
TA	Total area
TCA	Tricarboxylic acid
TCC	Transitional cell carcinoma
TSP	3-(trimethylsilyl)propionic acid
VIP	Variable importance to the projection
VOCs	Volatile organic compounds

1. Introduction to Metabolomics

1.1. Definition and metabolomics workflow

Metabolomics is the most recent “omic” science (1, 2) that arises as a complementary tool for genomics and proteomics. These two-last mentioned “omic” sciences are not yet fully understood and none of them considers the cellular activity at a metabolic level, a fact that is overcome through metabolomics. Therefore, metabolomics is the study of biochemical and biological processes that involve metabolites, the end products derived from cellular activity, which constitute the metabolome of an organism (3). Basically, it studies metabolite profiles consisting in unique and specific chemical features that represent the last biological reaction or response to a certain stimulus (4, 5). Metabolomics relies on the fact that any alteration in the normal functioning of systems biology caused by, for instance, the development of certain pathologies, affects the system homeostasis. Hence, metabolomics can provide detailed information on the physiological status of an organism, organ, cell or even a subcellular compartment (for instance, the mitochondria) (4), whether it is altered or not, through metabolite imbalances or stress (6). Metabolomics is also frequently named by a similar term, “metabonomics” (2, 7), which is not completely correct since metabonomics is seen as a more dynamic measure of metabolic responses to pathophysiological stimuli, genetic modification or response to diseases, drugs, toxins, and other adversities (8). On the other hand, metabolomics provides a comprehensive analysis of the entire metabolome through identification and quantification of all metabolites.

Metabolomic approaches can be classified as targeted or untargeted. In targeted metabolomics, specific well-known metabolites or a group of metabolites biochemically related are evaluated qualitatively and quantitatively (9, 10). On the contrary, untargeted metabolomics consists in the analysis and measurement of all possible metabolites present in a given biological system, and comparison between classes of samples (9), giving the opportunity to explore and discover new biochemical features and pathways. Independently of the approach, the main goal of metabolomic studies is the detection of a great number of metabolites without the requisite of exhaustive sample preparation (11).

Several studies have provided insight into how metabolomics can be useful in different fields of investigation (6, 10, 12-15), such as in the pharmaceutical and biotechnological fields (4, 7, 16, 17), and clinical medicine, particularly in areas such as biomarker discovery for disease diagnosis, prognosis, staging and treatment selection, drug development, monitoring therapies and evaluation of treatments effectiveness at earlier stages of a malignancy (4, 9, 18-24). From a therapeutic perspective, metabolomics is also useful for the development of biomarkers that provide information on

how different individuals will respond to drugs or to certain treatments (6). Particularly in cancer, metabolomics can be applied in the development of new diagnosis methods through the identification of biomarkers that might enable early diagnosis, monitoring of disease progression, prognosis, treatment guidance, drug discovery and development of new therapies (25, 26). The development of new biomarkers might overcome the limitations of classical diagnostic methods, since these often detect a malignancy in a rather late stage, where, in most cases, therapeutic possibilities and survival rates are lower (6). Furthermore, in cancer therapies, metabolomics can be applied in the monitoring of radiotherapy (27).

The potential of metabolomics and its several applications in oncology has increased. Indeed, metabolomics applied to cancer research has grown exponentially in the last decade, owing mainly to the discovery that cancer cells present an altered metabolism compared to normal cells (6, 28). In fact, it is well-known that cancer cells have greater energetic needs for growth, proliferation and survival, known as the "Warburg effect" (29, 30). Noteworthy, some studies have shown that oncogenes can interfere with metabolism (28, 31) and, in turn, alterations in the metabolism can cause mutation and activation of oncogenes (32). This highlights the importance of monitoring metabolic changes as it might help in both cancer detection and the comprehension of its mechanisms (33). Early diagnosis is one of the main goals of cancer metabolomics since in some cancers, such as bladder cancer, early detection and screening is still insufficient by the classical methods. Along with the development of metabolomics and its application in cancer research, a new concept has emerged: the cancer metabolome, which includes metabolites potentially suitable as biomarkers and particularly relevant to both oncologic processes and systemic responses to the malignancies (34).

Metabolomic studies have been performed in a wide range of samples, from cells *in vitro* (35-39) to tissues (20, 21, 40-43) or biofluids (22, 44-51), as well as hair (52), breath (15) and faeces (53, 54), which can be analysed through NMR or MS-techniques. The main steps of a metabolomic study include sample collection and preparation, data acquisition, statistical analysis and interpretation, as illustrated in figure 1. Each step will be described in more detail in the following topics.

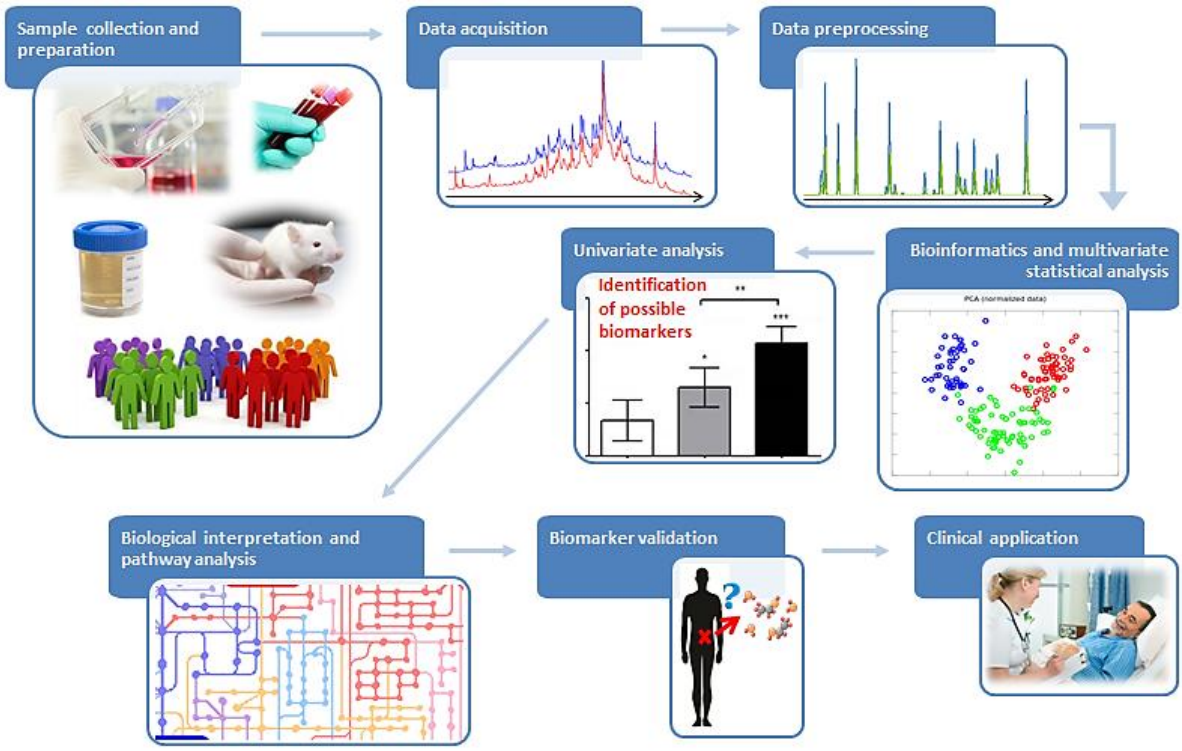

Figure 1. Schematic representation of the general workflow of metabolomics studies.

1.2. Samples Collection, Storage and Preparation

The most used biological model systems in metabolomic studies include *in vitro* cancer cultured cells, *ex vivo* cancer tissues and human biofluids, more commonly urine and serum (26). In any metabolomic study, the first step involves sample collection, storage and preparation (44), important procedures that will determine the integrity and successfulness of the metabolic analysis (55). Ideally, samples should be analysed without any preparation steps, since they might lead to the appearance of contaminants or some metabolites, or loss of others (56). Since that is not possible, it becomes crucial to ensure that samples are handled, stored and prepared appropriately and uniformly to avoid biased analysis and to allow the direct comparison between samples in terms of metabolites composition (10). Also, the collection and storage conditions are critical for protecting the metabolites from degradation or inhibiting the appearance of other foreign metabolites (for example, due to bacterial activity, especially in urine) (56, 57) that would contaminate the samples and render them unviable.

Factors such as age, gender, ethnic origin, diet, fasting state, life-style (smoking, drug or alcohol abuse), physical activity, diseases, and time of the day of sample collection should be taken into account in order to minimize their effects on the

metabolomic results (55). Therefore, it might be also important to collect information from the subjects/patients because it can aid in the interpretation of the results. Recruitment of best matched control populations is mandatory to reduce possible variations that arise from the factors mentioned above.

1.2.1. Collection and Storage of samples

Handling conditions depend on the type of matrix that will be analysed. Tissue samples collection require immediate inhibition of the metabolism to avoid the induction of "stress" metabolites (from hypoxia, for instance). The most common and sufficient method is to snap-freeze tissues in liquid nitrogen (55, 58).

As for serum or plasma, the first one is obtained by removing the natural clot from the blood, and the second one is the supernatant phase obtained by mixing blood with an anticoagulant, followed by a swift centrifugation at 4ºC (58). Lithium heparin is the most preferred anticoagulant in a metabolomic study since it does not interfere with the metabolic screening and profiling, unlike EDTA, for instance (58). After collection, samples must me immediately processed and stored in sterile containers at or below -70ºC (55, 56).

Regarding urine, its composition presents a diurnal variation that affects the time of collection (59), unlike other samples. As there is no standard method for urine collection and handling in metabolomics (44), it is of utmost importance to ensure that its collection and manipulation is uniform to avoid discrepancies. There are three different techniques for collecting urine samples: spot urine, which is collected at the time of a medical consultation (58); timed urine samples, used when studying time-related variations, especially in metabolites with high diurnal alteration (60); and 24h sample collection, which consists in collecting all urine produced within a day (61). The main disadvantage of spot urine is the variation in metabolites composition due to diet, circadian rhythm or lifestyle factors, which the 24h or first-pass urine collection may minimize (58). However, it has been reported that variations caused by pathologies are more significant that those caused by sample collection (62). It was also demonstrated that there is no significant difference, in terms of metabolic profile, between urine samples that are immediately frozen from those stored initially at 4ºC for 24h (63). Other studies have shown that urine samples are stable up to 6 months, either at -20 or -80ºC, and after multiple freeze-thaw cycles (until nine cycles) (62, 64). The method of freeze-drying has proven to cause instability of the urinary metabolites (64). Moreover, it was demonstrated that it is not obligatory to add a preservative in urine samples stored at below -20ºC (64).

Another characteristic of urine is the high abundance in urea that can hinder the detection of metabolites that co-elute with urea intense peak (65). Adding urease to urine samples to remove urea seemed to be a solution; however, it was shown that urease can also interfere with the identification of some metabolites (66). To overcome this problem, there is the possibility of optimizing the chromatographic parameters, for instance, so that co-elution is withdrawn (65). Bacterial activity is also a drawback when working with urine, being common the addition of sodium azide to inhibit it, without interfering with the stability of the samples (59).

Lastly, cell culture model systems, which are divided into primary cell culture, cell lines and cell strains, present a different collection approach, which depends on whether they grow adherently to an inert surface or in suspension in a liquid medium (67). In the case of suspension culture, cells can be collected from the medium through centrifugation (68) or filtration (69). Subsequently, to obtain the metabolic profile, quenching (responsible for inhibiting all cellular activities that might contribute to metabolic changes (67)) and extraction protocols are applied. The most recommended quenching technique is the use of cold (0.5ºC) isotonic saline (0.9% w/v NaCl), since it proved to be non-destructive and effective in inhibiting ADP and AMP generation from ATP (70). As for the extraction step, it was shown that using cold 50% aqueous acetonitrile provided a greater metabolite recovery and concentration (70). In turn, adherent cells require an additional first step before quenching and metabolite extraction, namely their harvesting from the culture flasks that can be achieved by trypsinisation or cell scrapping (67). Nevertheless, trypsin method has proven to be less suitable for *in vitro* metabolomics since it might cause membrane damage and metabolite loss, and consequently, lead to poorer metabolite content (71). Therefore, cell scrapping seems to be the best method, particularly when performed in an extraction solvent so that cell quenching is simultaneous. Among the possible extraction solvents, cold 80% methanol showed the highest extraction efficiency and reproducibility (72).

Independently of the type of culture (suspension or adherent cells), the standardization of culture conditions is crucial to better control the cell growth rate and phenotype stability so that reproducibility of the metabolomic data is higher (73). Indeed, any discrepancies in the number of harvested cells will lead to incorrect interpretation of the metabolomic results. Another important factor is the culture medium because it determines the cell growth and any variation in its composition will result in different consumption rates of substrates, affecting the cells metabolic profile (70, 73). Hence, it is recommended the use of the same medium in all experiments to avoid variations, even when dealing with different cell types. Nevertheless, the use of the same culture medium

Another characteristic of urine is the high abundance in urea that can hinder the detection of metabolites that co-elute with urea intense peak (65). Adding urease to urine samples to remove urea seemed to be a solution; however, it was shown that urease can also interfere with the identification of some metabolites (66). To overcome this problem, there is the possibility of optimizing the chromatographic parameters, for instance, so that co-elution is withdrawn (65). Bacterial activity is also a drawback when working with urine, being common the addition of sodium azide to inhibit it, without interfering with the stability of the samples (59).

Lastly, cell culture model systems, which are divided into primary cell culture, cell lines and cell strains, present a different collection approach, which depends on whether they grow adherently to an inert surface or in suspension in a liquid medium (67). In the case of suspension culture, cells can be collected from the medium through centrifugation (68) or filtration (69). Subsequently, to obtain the metabolic profile, quenching (responsible for inhibiting all cellular activities that might contribute to metabolic changes (67)) and extraction protocols are applied. The most recommended quenching technique is the use of cold (0.5ºC) isotonic saline (0.9% w/v NaCl), since it proved to be non-destructive and effective in inhibiting ADP and AMP generation from ATP (70). As for the extraction step, it was shown that using cold 50% aqueous acetonitrile provided a greater metabolite recovery and concentration (70). In turn, adherent cells require an additional first step before quenching and metabolite extraction, namely their harvesting from the culture flasks that can be achieved by trypsinisation or cell scrapping (67). Nevertheless, trypsin method has proven to be less suitable for *in vitro* metabolomics since it might cause membrane damage and metabolite loss, and consequently, lead to poorer metabolite content (71). Therefore, cell scrapping seems to be the best method, particularly when performed in an extraction solvent so that cell quenching is simultaneous. Among the possible extraction solvents, cold 80% methanol showed the highest extraction efficiency and reproducibility (72).

Independently of the type of culture (suspension or adherent cells), the standardization of culture conditions is crucial to better control the cell growth rate and phenotype stability so that reproducibility of the metabolomic data is higher (73). Indeed, any discrepancies in the number of harvested cells will lead to incorrect interpretation of the metabolomic results. Another important factor is the culture medium because it determines the cell growth and any variation in its composition will result in different consumption rates of substrates, affecting the cells metabolic profile (70, 73). Hence, it is recommended the use of the same medium in all experiments to avoid variations, even when dealing with different cell types. Nevertheless, the use of the same culture medium

is not always possible, when different cell lines only grow in different optimal medium. Otherwise, cells would grow inadequately or die (73).

For all sample types, it is recommended the storage in multiple aliquots *post* collection to avoid sample loss and degradation from various freeze/thaw cycles in a multiple metabolomic analysis (58, 63).

1.2.2. Sample treatment

Sample preparation involves several steps, such as extractions or use of buffers, which must be consistent and standardized as much as possible (55). Methods such as protein precipitation or metabolite extraction are widely used to reduce the interference of unwanted compounds (56). Nevertheless, there is not a protocol that solely covers the total metabolome (56). Therefore, it is becoming more common the performance of different extraction methods that, when combined, result in a larger number of metabolites.

1.3. Data acquisition

1.3.1. Analytical techniques

There are different analytical platforms that can be employed in a metabolomic study, each one presenting their advantages and limitations. When choosing which equipment to use, several factors must be considered, namely the aims of the study, the nature of the samples and the compounds that will be analysed, and also the financial capacity of the laboratory (56). The most widely employed analytical tools are nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS)–based techniques (2, 56). Once again, there is not a single platform that covers the entire metabolome, which is why it is desirable to combine them in order to obtain more information.

1.3.1.1. Separation techniques

In metabolomics, the separation of the analytes is generally by chromatographic methods, although electrophoresis (especially capillary electrophoresis) can also be applied (2).

The most widely used technique is gas chromatography (GC), which is usually coupled with mass spectrometry (GC-MS). GC advantage is its high chromatographic resolution; however, only volatile compounds (low molecular weight compounds) can be

analysed or those that can be volatilized. Other biomolecules, that are non-volatile and thermally labile, cannot be analysed unless they go through derivatization processes (2, 56). On the other hand, liquid chromatography (LC), also commonly coupled with MS, or high performance liquid chromatography (HPLC) have a lower chromatographic resolution and matrix effects are still challenging but, unlike GC, they can be used for a wider range of compounds (2), requiring lower amounts of sample, and derivatization steps are not needed (58).

1.3.1.2. Detection techniques

NMR provides detailed information on the structure of the analytes (56); it is non-discriminating, since it can detect all compounds with a resonating nuclei (e.g., ^{1}H, ^{13}C, ^{15}N) and non-destructive, since sample preparation is easy and minimal (6, 26, 56), allowing the recovery of sample for other analysis (2). Besides, it can analyse either liquid or solid samples, such as integral tissues or cell extracts (11, 26), presenting high reproducibility and resolution (6, 26, 56). The main disadvantage of NMR is its relatively lower sensitivity compared to MS (2, 74) and requires large amounts of sample (56). Other limitations include more expensive acquisition and maintenance, the special training needed to use NMR and the fact that it is not commonly used in clinics (6, 74).

The MS-techniques are used to identify and to quantify analytes after their separation by GC or LC (2), despite quantification is more difficult in MS than in NMR (6). MS is both sensitive and very specific (2, 6), and the separation of complex mixtures is more efficient (56). It has a high resolution, but the reproducibility is low, especially for biofluids (6). Additionally, it is more widely used in clinics than NMR (74). The major drawback of MS-techniques is the time-consuming sample preparation before analysis (6, 26, 56), such as derivatization, that can lead to sample alteration and, consequently, to inaccurate conclusions.

1.3.2. Quality control (QC)

In any metabolomic study, it is important to include quality control (QC) samples to ensure that the detection of metabolite differences among samples is not hindered by undesired sources of instrumental deviations, such as contamination peaks (2, 73). When planning a metabolomic experiment, it should be included enough QC samples from each population, so that it is possible to achieve metabolites validation and to minimize the probability of false results. Thus, the purposes of these samples, together with

standardization, is to ensure and optimize the reliability and reproducibility of the data obtained during metabolomic analysis, to determine the analytical precision and provide the possibility of performing signal correction so that signal variations are minimal (73, 75). QCs are also useful to evaluate the quality of samples, whether they are degraded or contaminated, which will affect the results (59). There are four main types of QCs that can be employed in a metabolomic study (55):

a) Chemical shift standards, namely 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) or 3-(trimethylsilyl)propionic acid (TSP), or pH standards (e.g., imidazole) for NMR analysis; in case of quantitative MS analysis, it is recommended leucine-enkephalin or labeled reference chemical standards;
b) Synthetic sample of around 30 to 40 representative compounds in biofluids or tissue metabolomic analysis, also useful in intra-lab quality control;
c) Pooled samples from the study, useful to correct batch effects, particularly in studies with high volume of samples;
d) Pooled human blood and urine standard reference materials (SMR) from the National Institute of Standards and Technology (NIST), which is available to all laboratories worldwide, being useful for inter-lab quality control.

1.4. Bioinformatics and statistical analysis

When performing an untargeted metabolomic study, rather than a targeted approach, the amount of data is immense, making difficult to analyse it without specialized software. After collecting the data, and before the statistical analysis, NMR or MS spectra need to be corrected to lessen the variations caused by experimental factors (56), such as those inherent to the analytical platforms (e.g., day-to-day alterations in sensitivity) or during sample preparation. Thus, metabolomics data need to be processed and normalized, a step that includes peak normalization and alignment, baseline correction, noise reduction, and deconvolution of peaks (important in cases of co-elution of compounds) (10, 58). Nowadays, these processes can be achieved more rapidly owing to the several bioinformatics software available (MZMine, XCMS, and others) (56, 58). Furthermore, in *in vitro* metabolomic studies, some additional and particular normalization methods are used to determine the cellular density, because it affects the metabolite concentration. Only then it is possible to compare the metabolite profile among samples of highly proliferative cells (67, 73). These methods include cell counting, protein or DNA

quantification, total metabolite signal or specific metabolic markers, being the DNA quantification the most robust strategy due to its higher accuracy and consistency of the measurements (73, 76).

After data pre-processing, which yields a great number of challenging variables to interpret (58), statistical analysis takes place. Multivariate statistical techniques can be classified into unsupervised or supervised. In the unsupervised approaches, no information is provided regarding samples classification so that they are separated according to their metabolite composition, leading to a prompt identification of possible similarities or differences between samples and also to the detection of outliers (11, 77). Principal component analysis (PCA) is the most commonly used unsupervised method and usually performed first. Subsequently, supervised approaches are applied, such as partial least squares-discriminate analysis (PLS-DA) or orthogonal PLS-DA (OPLS-DA) (77). In these methods, prior information on sample classes is given and they are mainly used to uncover the discriminative features between different classes of samples (26, 77). Once the potential metabolic signatures are identified as discriminative biomarkers, the metabolites must be identified. This part of a metabolomic study can be challenging since most spectral peaks can be unknown. Nevertheless, several metabolomics databases (2, 56, 58), such as the Human Metabolome Database (HMDB) or METLIN, can be helpful when identifying the peaks. However, for a more accurate identification, it is recommended to use standards of the respective analytes (2).

After the identification of potential metabolomic biomarkers, it is important to evaluate their functional and biological relevance with, for instance, metabolic pathway analysis. There are also tools available that provide remarkable representations of countless pathways (58), such as the Kyoto Encyclopaedia of Genes and Genomes (KEGG) database. Finally, additional studies must be performed to test and validate the potential metabolites (2) so that they become relevant as diagnostic or prognostic biomarkers. Validation studies are, generally, performed using larger and biologically diversified independent sample sets that highly represent the human population of interest. The main purposes of these studies are to check whether a biomarker or panel of biomarkers are associated with a disease and how accurate and specific on they are, so that its application in clinical research is enabled (78-80).

1.5. General challenges and limitations of metabolomics

The full employment of metabolomics in clinical practice is still challenging, despite the progresses and efforts in the search of biomarkers that could improve early diagnosis of diseases, especially in cancer.

Firstly, it is difficult to deal with the complexity and heterogeneity of the human metabolome, because of the high number of classes of compounds (peptides, amino acids, carbohydrates, organic acids), and the wide range of interactions and biological pathways in which they are involved (81). It would be necessary the combination of several analytical platforms to cover the whole metabolome (81). Still, even with only one analytical tool, the amount of data generated is great and complex enough to make it difficult to interpret without the application of very powerful bioinformatic tools (11, 56), as described previously. Likewise, there is the complexity and heterogeneity inherent to samples (whether they are biofluids or tissues) and patients, which are associated with several confounding factors such as age, gender, genetic background, lifestyle factors, diet, presence of other pathological conditions and respective therapies, surgical intervention (58, 82), and even metabolites derived from bacterial metabolism, a particular problem in urine samples (56). Consequently, this leads to "metabolic variability" (83) and false results, highlighting the importance of eliminating the confounding factors as much as possible. Hence, to overcome those confounding factors it is essential the inclusion of, besides healthy controls (HCs), symptomatic controls (SCs) (84), which consist of a group of patients that manifest the same symptoms as those with the disease of interest, but they do not present any clinical findings of that disease whatsoever, at the time of sampling (85).

Furthermore, in different studies that focus on the same disease (for example, bladder cancer) aiming at the discovery of novel diagnosis biomarkers, there is relatively little overlap between the identified metabolites, even when the same specimens were analysed (21, 22, 41, 46, 86-90). These inconsistencies are likely due to diverse sample collection and manipulation, work environment conditions, different planning and study design and the use of different analytical techniques. Thus, it is evident the lack of standardization methods when performing a metabolomic study. Besides, it is evident the insufficiency of the sample's size (either of disease or healthy controls) among studies as well, which lessens the reliability and the importance of the results.

Moreover, most studies do not test if metabolites found result from systemic responses to cancer or are actually produced by cancer cells. This is particularly important in *in vitro* studies, where the cancer-host interactions are lost (91, 92), as well as the

environment created within and around the tumour. These might explain the discrepancies of the results observed in *in vitro* studies and studies using tissues or biofluids. Besides, the confounding factors, that cause metabolomic variability, do not affect *in vitro* studies, which may be considered a benefit or a drawback, especially in the translatability to clinics since genetic and biological factors can influence the diseases outcome and progression, particularly cancer. Thus, although *in vitro* models are not as complex as the other matrices and show potential in helping the discovery of putative biomarkers, the interpretation of the results must be cautious and rigorous. The same applies when dealing with animal model systems, because of the inter-species genetic and metabolic differences, or human samples, due to inter- and intra-variability. In fact, a multiple matrix metabolomics study using human tissue and biofluids would be interesting to improve the comparison of the results and unveil which sample type is a better proxy of a certain disease.

It is crucial to carefully reflect on all those factors, because the interpretation of validation assays depends on it. Evaluation and validation studies are the most difficult step in a metabolomic experiment, besides being expensive, and time and sample consuming; however, they are necessary to enable the translation of the biomarkers found into clinical practice (83).

In cancer metabolomic studies, another setback must be taken into account, which is the fact that most metabolic alterations are common among various cancer types. Therefore, in these cases, validation studies of potential biomarkers should include other malignancies to better evaluate their specificity. Still concerning this setback, overcoming it might go through the determination and consideration of a panel of biomarkers rather than a single one (93-96).

1.6. Particularities of in vitro metabolomics for VOCs analysis

Volatile organic compounds (VOCs) are carbon-based molecules that present high vapour pressure at room temperature. They are classified according to their molecular weight and boiling point (50ºC - 260ºC) (97). Organisms can produce and release these compounds in exhaled breath or through body fluids such as urine, blood, sweat or faeces (53, 98-100), which in turn can be collected from the headspace (HS) of those matrices. In the case of *in vitro* studies, VOCs can also be collected from the HS of cells' extracellular medium (39).

The curiosity to investigate volatiles and their potential as diagnosis biomarkers has started with the observation that distinct diseases can be characterized by urine, exhaled breath or sweat specific odours (101), which are correlated with different VOCs patterns. Particularly for cancer, it was demonstrated that dogs can be trained to diagnose bladder (102), colorectal (103), lung and breast cancers (104) and melanoma (105) by sensing patients, supporting the idea that different diseases do have a characteristic “smellprint” (106). Therefore, there has been a great interest to study the cancer volatilome and a great attempt to discover VOCs as potential biomarkers for disease diagnosis (39, 53, 98, 103, 107-111). This new approach relies on the fact that VOCs are final products of cellular metabolism (106) and they are released through cell membranes, such as phospholipids or carbohydrates. Consumption or release of specific VOCs depends on cell membranes composition, which vary according to gene and/or protein characteristics (112), or if affected by oxidative stress and lipid peroxidation that cause membrane modification (106). Evidently, tumour cells will emanate a distinct pattern of VOCs due to their genetic alterations or protein modifications that affect oxidative stress and consequent peroxidation of membrane components (106).

The analysis of VOCs presents several advantages including simple and fast sample preparation, compared to other molecules that need a derivatization process (e.g. amino acids and fatty acids), which prevents loss of metabolites; non-invasive collection (easily detected in exhaled breath or urine); enables high-throughput screening and analysis of numerous compounds across different samples (113). Considering *in vitro* VOCs collection, this approach may help to understand the source of VOCs as cell metabolomics present more controllable experiment designs and easier interpretation of the results due to the absence of factors related to age, gender and other individual differences that affect the metabolic profile (73).

When performing an *in vitro* metabolomic study, the whole cell metabolome can be investigated or instead, either the endometabolome, which comprises all metabolites present inside the cell, or the exometabolome, which is represented by all metabolites in the cell extracellular culture medium (73, 97). The exometabolome profile results from the interchange of compounds between cells and the culture medium. This will reflect cells metabolism, which differs if variations in the culture medium or in the experimental design are introduced (73). The study of the exometabolome is highly favoured over the endometabolome in VOC-biomarker investigations due to the fact that endometabolomic approaches require additional cell membranes disruption and concentration of the collected compounds, which can lead to a substantial loss of VOCs (73).

In *in vitro* studies, two main matrices are used for VOCs collection, namely the headspace (HS) of cell-free culture medium or HS of medium containing cells (97). These two approaches present some differences in the extraction procedures, such as the temperature of the analysis, which is 37 ºC in the presence of cells but higher in cell-free culture medium; the addition of salts or alteration of the pH in cell-free culture medium to improve the efficiency of the analysis, which cannot be employed in the presence of cells. In contrast, analysing medium samples with cells avoids the loss of VOCs during storage (97, 114). Regarding the techniques used to collect VOCs, some enhancements have been made over the years to improve sample preparation and VOCs extraction methodologies. Among the possible techniques that can be employed, solid-phase microextraction (SPME) is one of the best options as it provides a simple, economical, solvent-free, low time-consuming method for VOC extraction, sample pre-concentration and analysis, enabling minimal sample treatment and, consequently, minimal alteration of sample metabolites (115, 116). An analogous technique is headspace-SPME (HS-SPME), in which compounds present in the gas space are analysed, and their adsorption to the SPME fibre depends on the equilibrium between the sample and the HS, as well as between the HS and the fibre coating. The efficiency of HS-SPME depends on several factors including the chemical nature of the compounds (volatiles or easily volatilized), the type of fibre coating, extraction temperature and time, salting-out effect and the type of sample (whether it is biological, or in solid, liquid or gas phase) (117, 118). Nowadays, it is widely used HS-SPME coupled with GC-MS because it allows a rapid, efficient and automated VOCs extraction followed by immediate identification.

The limitations related to general metabolomic studies in terms of lack of standardized extraction and analytical methods also include *in vitro* VOCs studies, which hampers the possibility to compare results among different studies. Other aspects of *in vitro* approaches that contribute for the discrepancies found among cell metabolomic studies include the use of different cell types and cell culture medium, different cell controls, different cell density and period of cultivation and growth, and different methodology for collection, handling and storage of culture medium samples (97). Another possible limitation of *in vitro* VOC studies is that the metabolites found altered in cancer cell lines and considered potential biomarkers may not be found in *in vivo or ex vivo* analysis. It has been proposed that hyperoxic cell culture conditions lead to those divergences since cancer tissues are associated with hypoxic (low oxygen) or anoxic (absence of oxygen) growth conditions, in contrary to normal tissues (119). Different cellular oxidative status leads to different VOC metabolism. Thus, *in vitro* studies in

hypoxic or anoxic conditions would be a worthy approach in order to better reproduce tumours environment.

Other factors can hamper the translation of *in vitro* to *in vivo* or *ex vivo* samples. For instance, VOCs collected from culture cells may not be endogenous and rather be originated from other sources (culture flasks, extraction procedures and sampling conditions) (114, 120). In addition, different experimental design (from sample collection to statistical analysis techniques), variations related to patients' genetic and pathological characteristics and the low number of *in vitro* studies compared with other matrices also contribute for the discrepancies between the matrices. Furthermore, very little is known about VOCs metabolism and how they are produced by cancer cells or what they are consumed for. This is in fact considered the main challenge of VOCs analysis in cancer biomarker discovery (97).

Despite the expectable low reproducibility of VOCs translation from *in vitro* to other biological matrices, some studies have, however, shown common VOCs altered in cancer cells and tissues, urine or exhaled breath (39, 107, 121, 122). This encourages furthering investigation of potential VOC-biomarkers that can discriminate several cancer types as well as other diseases, so that they can be employed as a diagnosis tool in clinics.

2. METABOLOMICS RESEARCH IN BLADDER CANCER

Review Paper

Biomarkers in bladder cancer: A metabolomic approach using *in vitro* and *ex vivo* model systems

Reprinted from International Journal of Cancer, 139: 256-268

IJC
International Journal of Cancer

Biomarkers in bladder cancer: A metabolomic approach using *in vitro* and *ex vivo* model systems

Daniela Rodrigues[1], Carmen Jerónimo[2,3], Rui Henrique[2,3,4], Luís Belo[5], Maria de Lourdes Bastos[1], Paula Guedes de Pinho[1] and Márcia Carvalho[1,6]

[1] UCIBIO/REQUIMTE, Laboratory of Toxicology, Department of Biological Sciences, Faculty of Pharmacy, University of Porto, Porto, Portugal
[2] Cancer Biology & Epigenetics Group, Portuguese Oncology Institute-Porto, Porto, Portugal
[3] Department of Pathology and Molecular Immunology-Biomedical Sciences Institute Abel Salazar (ICBAS), University of Porto, Porto, Portugal
[4] Department of Pathology, Portuguese Oncology Institute-Porto, Porto, Portugal
[5] UCIBIO/REQUIMTE, Laboratory of Biochemistry, Department of Biological Sciences, Faculty of Pharmacy, University of Porto, Porto, Portugal
[6] FP-ENAS, CEBIMED, Fundação Ensino e Cultura Fernando Pessoa, Universidade Fernando Pessoa, Porto, Portugal

Mini Review

Metabolomics has recently proved to be useful in the area of biomarker discovery for cancers in which early diagnostic and prognostic biomarkers are urgently needed, as is the case of bladder cancer (BC). This article presents a comprehensive review of the literature on the metabolomic studies on BC, highlighting metabolic pathways perturbed in this disease and the altered metabolites as potential biomarkers for BC detection. Current disease model systems used in the study of BC metabolome include *in vitro*-cultured cancer cells, *ex vivo* neoplastic bladder tissues and biological fluids, mainly urine but also blood serum/plasma, from BC patients. The major advantages and drawbacks of each model system are discussed. Based on available data, it seems that BC metabolic signature is mainly characterized by alterations in metabolites related to energy metabolic pathways, particularly glycolysis, amino acid and fatty acid metabolism, known to be crucial for cell proliferation, as well as glutathione metabolism, known to be determinant in maintaining cellular redox balance. In addition, purine and pyrimidine metabolism as well as carnitine species were found to be altered in BC. Finally, it is emphasized that, despite the progress made in respect to novel biomarkers for BC diagnosis, there are still some challenges and limitations that should be addressed in future metabolomic studies to ensure their translatability to clinical practice.

Bladder Cancer and the Need for Novel Diagnosis Strategies

Bladder cancer (BC) is the second most common human malignancy affecting the urinary system and one of the most deadly cancers worldwide.[1] Its incidence continues to rise and mortality rates have remained unchanged over the past 3 decades, since successful treatments depend mainly on early detection.[2] Risk factors are associated with environmental, dietary/lifestyle, especially smoking and genetic factors.[3,4]

BC is a heterogeneous and multifocal malignancy, being divided in 3 main histological subtypes. Most BCs are transitional cell carcinomas (TCC) (90%),[5] also called urothelial cell carcinomas since they develop from the cells of the bladder lining (urothelium). The other two types, squamous cell carcinoma and adenocarcinomas, represent 10% of BC.[5] Noteworthy, metabolomic studies on BC are, however, mostly focused on urothelial bladder carcinoma. BC can be classified as low-grade or high-grade, depending on the degree by which cancer cells histologically differ from normal bladder cells, being high-grade BC more aggressive and invasive than low-grade.[6] Furthermore, BC is also classified as superficial or nonmuscle invasive bladder cancer (NMIBC) and invasive or muscle invasive bladder cancer (MIBC), based on the level of invasion of the muscular bladder wall.[6] Whereas NMIBC is restricted to the lining of the bladder, MIBC spreads and invades the muscle wall and other tissues/organs, and might also cause metastatic spread.[6,7] Remarkably, these two tumor types display quite distinct molecular characteristics. Indeed, NMIBCs are commonly near-diploid and have fewer genomic rearrangements than MIBCs, which are often aneuploidy. Moreover, up to 80% of NMIBC have activating point mutations in FGFR3, whereas MIBC only displays this alteration

Key words: bladder cancer, biomarkers, metabolomics, metabolic pathways, review, cell lines, tissue, urine, serum, biofluids
Grant sponsor: National Funds from FCT (Fundação para a Ciência e Tecnologia) and FEDER under Program PT2020; **Grant number:** Project 007265-UID/QUI/50006/2013.
DOI: 10.1002/ijc.30016
History: Received 3 Nov 2015; Accepted 19 Jan 2016; Online 25 Jan 2016
Correspondence to: Daniela Rodrigues; FP-ENAS, CEBIMED, Fundação Ensino e Cultura Fernando Pessoa, Praça Nove de Abril, 349 4249-004 Porto, Portugal, Tel.: +351 225071300; Fax: +351 225074630; E-mail: daniela.fgrodrigues@gmail.com or Márcia Carvalho; FP-ENAS, CEBIMED, Fundação Ensino e Cultura Fernando Pessoa, Praça Nove de Abril, 349 4249-004 Porto, Portugal, Tel.: +351 225071300; Fax: +351 225074630; E-mail: mcarv@ufp.edu.pt

in about 20% of cases. Conversely, EGFR expression is upregulated in only ~20% NMIBC, whereas this protein tyrosine kinase is overexpressed in about 50% of MIBC. Nevertheless, this categorization of BC failed to provided relevant prognostic or predictive information for clinical application and recent gene expression profiling studies found genetic signatures that cross grade and stage groupings, with superior prediction of clinical behavior.[8]

BC is mostly an asymptomatic disease at its inception, with no specific symptoms, hindering its detection. The most common symptom of BC is hematuria, occurring in 80%–90% of the cases[9]—however, it may be caused by other urinary diseases. Current standard diagnostic methods include cystoscopy and urine cytology. The first consists on the direct observation of the bladder but it fails to visualize certain regions and some cases of carcinoma *in situ*,[10,11] apart from being invasive and uncomfortable for patients. The latter consists in microscopical examination of urine smears to check for abnormal cells, blood and other markers.[12] Urinary cytology is noninvasive and highly specific, yet it lacks sensitivity to detect low-grade tumors.[13] Interestingly, there are already some validated and approved molecular markers, such as ImmunoCyt, UroVysion and bladder tumor antigen, among others.[14,15] Still, they lack specificity and its high costs limit their use in clinical practice.

Early detection of BC is pivotal to improve overall survival rate of patients. Since current diagnostic strategies are limited in detecting early BC, there is a great need to search for early, specific and cost-effective biomarkers, as well as noninvasive detection tools, to improve BC diagnosis. Novel BC biomarkers can also be valuable for prognostication, patient stratification and identification of patients at high risk for cancer recurrence and progression. As a result, better management of BC and successful treatment might be achieved.

Metabolomics, defined as the comprehensive analysis of all metabolites in a biological system, has gained importance over the last years in the quest for promising, alternative strategies for early cancer diagnosis. This "omic" approach has been applied in cancer research since altered metabolism is a widely recognized cancer hallmark.[16] Cancer cells have greater energetic needs than normal cells to fuel the high DNA replication, protein synthesis rates and fast, constant growth even in hostile environments. By reprograming their metabolism, cancer cells are able to buildup sufficient biomass and to adapt, survive and proliferate under conditions of stress.[17,18] Consequently, metabolites excreted by cancer cells will reflect alterations in the normal activity of metabolic pathways, particularly glycolysis, amino acid and fatty acid catabolism and its physiological status.[19]

Several studies have provided insight into how metabolomics can be useful in cancer research, showing its great potential not only toward the identification of candidate cancer biomarkers but also in a better understanding of the mechanisms underlying the pathogenesis of cancer. Metabolomics studies have been applied to several cancer types including those of the breast,[20,21] ovary,[22] prostate,[23,24] kidney,[25,26] lung,[27] brain[28,29] and liver,[30,31] with encouraging results. Concerning BC, the applicability of a metabolomics approach is limited compared with other cancers. Nevertheless, there are already some metabolomic studies on BC, which will be highlighted in this review, as well as identified potential biomarkers for BC early diagnosis.

Disease Model Systems for BC Metabolomics Studies

Current disease model systems used in the study of BC metabolome include *in vitro* cancer cell cultures, *ex vivo* neoplastic bladder tissues and human biofluids.[32] Each model system has specific advantages and drawbacks, most notably related to sample collection and handle and degree of complexity.

In vitro culture systems, which comprise immortalized cell lines and primary cell cultures, represent the least complex of the various disease models and has a number of advantages over tissue or biofluid analysis, including simpler and controllable experimental settings, reduction in animal testing and providing a better insight into the metabolic changes.[19] However, *in vitro* metabolomics has challenges related to extrapolation to *in vivo* systems.[33]

Ex vivo BC tumor tissues are also important to consider in these studies since it gives information on metabolites altered within the solid tumor and in its microenvironment.[34] Another advantage of using tissues is that the possibility of collecting cancer and matched normal tissue samples from the same patient strongly diminishes the influence of confounding factors. Nonetheless, tissue collection through surgical procedures is extremely invasive,[34] needs special equipment and expertise and requires inhibition of metabolism at the moment of collection to avoid the induction of "stress" metabolites.[35] Tissues also have a limited availability and high heterogeneity, making the samples preparation more difficult. Additionally, samples may be contaminated by surrounding cells, thus special techniques are required to heighten sample purity.[32,34]

Biofluids have the advantage of being the easiest samples to work with, being urine the most common sample analyzed in BC metabolomics.[34] Urine seems more likely to reflect metabolic alterations occurring in BC since it directly contacts with the tumor tissue and may contain tumor cells and metabolites shed from the bladder tumor.[36] The use of urine as sample for metabolomic studies has a number of advantages as compared with serum which include its noninvasive nature, easier collection and handle, has relatively high thermodynamic stability, needs less sample preparation, has higher amounts of metabolites, and it is a less complex matrix, in terms of protein content. Nevertheless, urine composition is highly affected by several confounding factors, such as age, environmental and dietary/lifestyle factors and waste products, which consequently affects the metabolomic profile,[15,37] and it is more prone to microbial contamination.

Moreover, as a significant number of patients will have the primary tumor removed (*e.g.*, cystectomy, nephroureterectomy), if these patients relapse, urine will no longer be in contact with the tumor and, accordingly, it is unlikely to be a good model system in these patients. Blood serum/plasma is also an important sample because its composition directly reflects the metabolic processes occurring in the organism, it is less affected by diurnal variations and confounding factors compared with urine.[34] However, its composition is more complex than that of urine.[34]

Noteworthy, the comprehensive study of metabolome is only possible through advanced and sophisticated methodologies. The analytical platforms most commonly used in metabolomics studies are proton magnetic resonance spectroscopy (^{1}H-NMR) and both liquid and gas chromatography (GC) coupled to mass spectrometry (MS).[35] Detailed description on these techniques, as well as the statistical tools for metabolomics data analysis, is beyond the scope of the this article; however, there are several studies on this matter.[18,32,35,38,39]

Metabolomic Studies in BC Model Systems and Identification of Potential Biomarkers

Significant progress in BC metabolomic research as well as a better understanding of metabolic alterations related to the development of this type of cancer has occurred. Studies on this matter report alterations in the levels of some metabolites in BC that could be potential candidate biomarkers for clinical practice, which are summarized in Tables 1 and 2.

Cell line studies

There are just a few *in vitro* metabolomics studies in the literature, all of them attempting at profiling the metabolome of immortalized BC cell lines (Table 1). The first study focused on the measurement of free and total glutathione (GSH) levels in 5 human BC cell lines, namely, UCRU-BL-13, -17 and -28, RT-4 and TCCSUP, and in 8 non-BC cell lines (used as controls).[40] It was observed that GSH levels were significantly higher in four of the five BC cell lines (UCRU-BL-13, -17 and -28, RT-4), with the average GSH concentration being 6-fold higher than in control cell lines.[40] Increased GSH concentrations, as well as overexpression of antioxidant enzymes,[59,60] are crucial for the survival of cancer cells under oxidative stress. Despite these results, GSH upregulation is a hallmark of cancer cells, meaning that it is not specific for BC. In fact, GSH levels were similarly ranked in other cancer types, namely, breast,[61,62] kidney,[63] liver,[61,64] brain,[65,66] lung,[67] bone and prostate.[61]

Pasikanti *et al.* (2010) carried out a metabolomic approach, using both GC-MS and GCxGC-TOF-MS techniques. Metabolic fingerprinting of immortalized nontumorigenic human bladder cells (HUC-1) and their neoplastic counterpart, HUC T-2 cells, revealed profound differences between their metabolic phenotypes. Five and 20 metabolites were identified by GC-MS and GCxGC-TOF-MS, respectively (Table 1). Among them, glycine was the metabolite found at higher levels, whereas 3-phosphoglyceric acid, a glycolytic intermediate, was at lower levels in the HUC T-2 cells media compared with that of HUC-1 cells.[41] This is concordant with the metabolic alterations observed in glycolysis since it shows that glycine production through 3-phosphoglyceric acid is upregulated. Furthermore, as glycine is an essential source of one-carbon units for nucleotide synthesis, it also emphasizes that this metabolic pathway is upregulated as well.

Dettmer *et al.* (2013) performed a study involving numerous immortalized human cancer cell lines, including BC cell lines J82 (high-grade) and RT4 (low-grade), to identify differences in their metabolic profiles. In total, 700 metabolite features were found, but only 42 were discriminative (Table 1). All cancer cells presented higher production of lactate, higher assimilation of amino acids and higher levels of extracellular arginine and nicotinamide.[42] This evidences the increase in biosynthetic activity and in bioenergetics demands of cancer cells.

To further investigate whether the metabolomic signature is altered with progression of cancer, a study was recently carried out to look at the glycolytic profile of RT4 and TCCSUP cell lines, low- and high-grade BC, respectively.[43] Samples collected from the extracellular medium were analyzed by ^{1}H-NMR. Glucose consumption was found to be similar in both cell lines, though it was slightly higher in the highly invasive cancer cells. TCCSUP cells also presented increased pyruvate consumption, higher levels of alanine and lactate than in the RT4 cells.[43] This suggests that alterations in the metabolism of pyruvate and in lactate production are somewhat related to BC development and progression, from lower to higher grades. In fact, the enhanced glucose uptake, required for ATP production, and the insufficient oxygen supply, caused by the rapid proliferation and deficient blood deliver which generates anaerobic conditions,[26] favors the conversion of pyruvate to lactate. In addition, the conversion of pyruvate to either lactate or alanine is coupled with the conversion of NADH to NAD^+, an essential metabolite for the maintenance of the glycolytic flux in cancer cells.[68]

Ex vivo tissue studies

Strengthening the importance of GSH in the protection against tumor microenvironment-related aggression, a study conducted by Pendyala *et al.* (1997) also showed higher levels of GSH in BC tissue compared with those of healthy controls.[40]

Putluri *et al.* (2011) performed a study aiming at the identification of potential candidate biomarkers through BC metabolomic signatures and unravel bioprocesses associated to BC carcinogenesis. The metabolic profiles of BC and benign-looking adjacent tissues, as well as healthy controls, were determined by LC-MS.[44] Fifty metabolites exhibited significant differences in BC tissues when compared with the benign ones. Among the perturbed metabolites, elevated levels of amino acids, namely, serine, asparagine, valine,

Table 1. Metabolic pathways and metabolites altered in human bladder cancer derived cell lines

Study [reference]	Cell lines: Controls	Cell lines: BC	Cell lines: Type of BC	Cell lines: Grade/ Stage	Analytical platform	Total/Distinct metabolites found	Discriminatory metabolites	Dysregulated metabolic pathways
Pendyala *et al.* 1997 [40]	8 (from carcinomas of ovary, colon, melanoma and glioblastoma)	UCRU-BL-13 UCRU-BL-17 UCRU-BL-28 RT4 TCCSUP	TCC TCC TCC TCP TCC	II III/T4 II/T4 I IV	HPLC/FLD	NA/1	GSH (+)	Glutathione metabolism
Pasikanti *et al.* 2010 [41]	HUC-1	HUC T-2	–	–	GC-MS GCxGC-TOF-MS	65/5 286/20	1. Gly (+); 2. Malate (+); 3. 1-Amino-cyclopropane-1-carboxylic acid (−); 4. 3-Methyl-2-oxo-butanoic acid (+); 5. Propanoic acid derivative (+); 6. Glycerol-2-phosphate (−); 7.Glucose(−); 8. 3-Phosphoglyceric acid (−); 9. Sugar alcohol (−); 10. Glucuronic acid (−); 11. Carbohydrate (+); 12. Sorbose (−); 13. Erythrose (−); 14. 1-Phenyl-ethylamine (−) 15. Malonic acid (+); 16. Benzyl alcohol (−); 17. Mannitol (−); 18. Myo-inositol (−); 19. FA derivative (+); 20.Unknown (−)	1, 3, 4. Amino acid metabolism; 2, 11. TCA cycle; 5. Pyruvate/ Amino acid metabolism; 6. TG metabolism; 7, 8. Glycolysis; 9, 12, 13, 17. Sugar alcohol metabolism; 10. Inositol phosphate metabolism; 14. Phenylalanine metabolism; 15. Pyrimidine metabolism/Alanine metabolism; 16. Aromatic compounds catabolism; 18. Phosphatidylinositol signaling system/Inositol phosphate metabolism; 19. FA metabolism
Dettmer *et al.* 2013 [42]	NS	J82 RT4	TCC TCP	III, IV/T3 I	LC-MS GC-MS	700/42	1. Lactate (+), pyruvate (−), glucose (−); 2. All amino acids (+) except Gly (−), Lys (−) and Cys (−); 3. MTA (+); 4. Citrulline (+); 5. Ornithine (+); 6. Glycerol (+); 7. Glycerate (−); 8. Fumarate (−), citrate (−); 9. Erythritol (+); 10. Fructose (+); 11. Myo-inositol (+); 12. Nicotinate (+), nicotinamide (+); 13. Kynurenine (+), indole-3-lactate (−), indole-3-acetate (−), hydroxyindole-3-acetate (−); 14. Serotonin (−); 15. Anthranilate (+); 16. Indole-3-propionate (−)	1. Glycolysis/Pyruvate metabolism/TCA cycle; 2, 3, 4, 5. Amino acid metabolism; 5. Glutathione metabolism; 6. TG metabolism; 7. Pentose phosphate pathway; 8. TCA cycle; 10. Sugar metabolism; 9. ABC transporter; 11. Phosphatidylinositol signaling system/Inositol phosphate metabolism; 12. Nicotinate/nicotinamide metabolism; 13, 14, 15. Tryptophan metabolism; 16. NA
Conde *et al.* 2015 [43]	-	RT4 TCCSUP	TCP TCC	I IV	^{1}H NMR	NS/4	1. Glucose (−); 2. Pyruvate (−); 3. Lactate (+); 4. Ala (+)	1, 2, 3. Glycolysis; 4. Amino acid metabolism/Pyruvate metabolism

Note: (+) indicates increased metabolite level in BC, (−) indicates decreased metabolite level in BC.
^{1}H NMR, proton nuclear magnetic resonance; Ala, alanine; BC, bladder cancer; Cys, cystine; FA, fatty acid; FLD, fluorescence detection; GC, gas-chromatography; Gly, glycine; GSH, glutathione (reduced form); HPLC, high-performance liquid chromatography; LC, liquid-chromatography; Lys, lysine; MS, mass spectrometry; MTA, 5'-deoxy-5'-methylthioadenosine; NA, not available; NS, not specified; TCA, tricarboxylic acid (cycle); TCC, transitional cell carcinoma; TCP, transitional cell papilloma; TG, triglycerides; TOF, time-of-flight.

Table 2. Clinical metabolomic studies of BC and identified potential biomarkers for early diagnosis

Sample type study [reference]	Study cohort size				Analytical platform	Total/Distinct metabolites found	Discriminatory metabolites	Dysregulated metabolic pathways
	Controls	BC	Age (years)	Gender				
Ex vivo bladder tissue								
Pendyala *et al.* 1997 [40]	*n*=37 (14 HC; 23 w/o cancer w/BC history)	*n*=17	NS	NS	HPLC/FLD	NA/1	GSH (+)	Glutathione metabolism
Putluri *et al.* 2011 [44]	*n*=25 (matched paired)	*n*=58 31 malign 27 benign	NS	NS	LC-MS	2019/50	1. Ser (+), Asn (+), Val (+), Lys (+), Ala (−), Phe (+), His (+), Ile (+), Leu (+), Tyr (+), Trp (+); 2. SAM (+); 3. Palmitic, lauric and oleic acids (−); 4. Creatine (+); 5. Carnitine (+), isobutyryl carnitine (+); 6. Taurine (−); 7. Coumarin (−); 8. Glyceraldehyde-3-phosphate (−); 9.Histamine (−); 10. Spermidine (+); 11.Aniline (+); 12. N1-acetylspermine (+), N8-acetylspermine (+); 13. Urocanic acid (+); 14. 2-Hydroxybutyric acid (−); 15. Homocysteine (−); 16. 4-Pyridoxic acid (−); 17. Pipecolic acid (+); 18. Aminobutyric acid (+); 19. Cytidine monophosphate (+); 20. Citramalic acid (+); 21. Niacinamide (+); 22. Hippuric acid (−); 23. Homoserine (+); 24. Thymine (+); 25. 5-Hydroxy indole acetic acid (−); 26. Hypoxanthine (+); 27. D-ribonolactone (+); 28. Kynurenine (+), 3-hydroxykynurenine (+); 29. 4-Hydroxy phenyl lactic acid (+); 30. Guanine (+); 31. Uracil (+); 32. Phthalic acid ester (−); 33. 1,7-dimethylxanthine; 34. Norepinephrine; 35. Pantothenic acid	1, 9, 10, 13, 15, 18, 20, 22, 23, 31, 34, 35. Amino acid metabolism; 2. Methylation; 3. Unsaturated FA metabolism; 4. Urea cycle; 5. FA metabolism; 6. Cysteine sulfinic acid pathway; 7, 33. Biosynthesis of secondary metabolites; 8. Glycolysis/Pentose phosphate pathway; 11. Drug metabolism; 12, 14, 17, 25, 27, 29. NA; 16. Vitamin B metabolism; 18, 21. Nicotinate and nicotinamide metabolism; 19, 24, 31. Pyrimidine metabolism; 26, 30. Purine metabolism; 28. Tryptophan metabolism; 32. Polycyclic aromatic hydrocarbon degradation
Tripathi *et al.* 2013 [45]	*n*=26 (benign)	*n*=33 17 LG 16 HG	LG:60.1 HG:68.8	29 M 4 F	^{1}H NMR GC-MS	NS/22	1. TG (−); 2. Leu (+), Ile (+), Val (+), Ala (+), Gly (+), Lys (+), Phe (+), Asp (+), Tyr (+); 3. Lactate (+); 4. Acetate (+); 5. Glu (+), Gln (+); 6. GSH (+); 7. Creatine (+); 8. ChoCC (+); 9. Myo-inositol (+); 10. Taurine (+); 11. UDP (+)	1. TG metabolism; 2, 5. Amino acid metabolism 3. Glycolysis; 4. Pyruvate metabolism; 5. Nucleotide synthesis; 6. Glutathione metabolism; 7. Urea cycle; 8. Glycerophospholipids metabolism; 9. Phosphatidylinositol signaling system/Inositol

Table 2. Clinical metabolomic studies of BC and identified potential biomarkers for early diagnosis (Continued)

Sample type study [reference]	Study cohort size				Analytical platform	Total/Distinct metabolites found	Discriminatory metabolites	Dysregulated metabolic pathways
	Controls	BC	Age (years)	Gender				
								phosphate metabolism; 10. Cysteine sulfinic acid pathway; 11. Pyrimidine metabolism
Biofluids: plasma or serum								
Lin *et al*. 2012 [46]	*n*=48 (20 HC, 10 prostate hyperplasia, 18 nephrolithiasis)	*n*=13	NS	NS	RPLC-MS HILIC-MS	NS/7	1. Eicosatrienol (+); 2. Azaprostanoic acid (+); 3. Retinol (+); 4. Docosatrienol (+); 5. Phosphatidylcholine (+); 6. Acetylphenylalanine (+), methyl hippuric acid (+)	1, 4. Metabolism of polyunsaturated FAs; 2. NA; 3. Retinol metabolism; 5. Glycerophospholipids metabolism; 6. Phenylalanine metabolism
Cao *et al*. 2012 [47]	*n*=73 (28 calculi patients, 25 HC, 20 post-TURBT)	*n*=39 17 LG 22 HG	NS	NS	^{1}H NMR	NS/11	1. Ile (−), Leu (−), Tyr (−), Phe (−), Gly (−); 4. Cho (−); 5. Lactate (−); 6. Citrate (−); 7. V/LDL (+); 8. Glucose (+); 9. Acetoacetate (+)	1. Amino acid metabolism; 4. Glycerophospholipids metabolism; 5, 8, 9. Glycolysis/pentose phosphate pathway; 6. TCA cycle; 7. Cholesterol metabolism; 9. Amino acid and Ketone bodies metabolism
Bansal *et al*. 2013 [48]	*n*=32	*n*=68 36 LG 32 HG	>40	All M	^{1}H NMR	22/6	1. DMA (+); 2. Malonate (+); 3. Lactate (+); 4. His (+), Val (+), Glu (+)	1. Glycerophospholipids metabolism; 2. TCA cycle/Pyrimidine metabolism; 3. Glycolysis/TCA cycle; 4. Amino acid metabolism/Nucleotide synthesis
Biofluids: urine								
Issaq *et al*. 2008 [49]	*n*=48 (HC)	*n*=41 28 LG 13 HG	M: 77.1 F: 74.6	36 M 5 F	HPLC-MS	NA/NA	NA	NA
Srivastava *et al*. 2010 [50]	*n*=70 (33 benign, 37 HC)	*n*=33	NS	NS	^{1}H NMR	NS/4	1. Taurine (+); 2. Citrate (−); 3. Phe (−), hippuric acid (−)	1. Cysteine sulfinic acid pathway; 2. TCA cycle; 3. Amino acid metabolism
Pasikanti *et al*. 2010 [51]	*n*=51 (HC)	*n*=24	67.2 ± 12.3	20 M 4 F	GC-TOF-MS	398/15	1. Val (+); 2. 2-butenedioic acid (−); 3. Citrate (−); 4. Glycerol (−); 5. Melibiose (+); 6. Ribitol (−); 7. Gluconic acid (−); 8. Uridine (+); 9. Senecioic acid (−); 10. 2,5-Furandicarboxylic acid (−); 11. Sumiki's acid (−); 12. 2-Propenoic acid (−); 13. Valerate (−); 14. Fructose (−); 15. Ribonic acid (−)	1, 13. Amino acid metabolism; 2, 3. TCA cycle; 4. TG metabolism; 5. Sugar metabolism; 6, 7. Pentose phosphate pathway; 8. Pyrimidine metabolism; 9, 10, 11, 15. NA; 12. Aromatic compounds catabolism; 14. Fructose/Amino sugar metabolism

Table 2. Clinical metabolomic studies of BC and identified potential biomarkers for early diagnosis (Continued)

Sample type study [reference]	Study cohort size				Analytical platform	Total/Distinct metabolites found	Discriminatory metabolites	Dysregulated metabolic pathways
	Controls	BC	Age (years)	Gender				
Huang *et al.* 2011 [2]	*n*=32 (HC)	*n*=27 21 LG 6 HG	56.0	19 M 8 F	HILIC-MS RPLC-MS	703/11 417/9	1. Carnitines C9:1, C8:1, C10:1 (−); 2. Acetylcarnitine (+), 2,6-dimethylheptanoyl carnitine (−); 3.Hippuric acid (−), fragment of hippuric acid (−); 4. Component I (+); 5. Trigollenine (−); 6. PAGN (−); 7. Leucylproline (−); 8. Phosphorylcholine (−), fragment of phosphorylcholine (−); 9. Unidentified (+)	1, 2. FA metabolism; 3. Amino acid metabolism; 4, 9. Unknown; 5. Nicotinate/nicotinamide metabolism; 6. Phenylalanine metabolism; 7. NA; 8. Glycerophospholipids metabolism
Putluri *et al.* 2011 [44]	*n*=51	*n*=83	NS	NS	LC-MS	35/25	NA	NA
Jobu *et al.* 2012 [52]	*n*=7 (HC)	*n*=9	NS	NS	GC-MS	12/5	1. Ethylbenzene (+); 2. Nonanoyl chloride (+); 3. Dodecanal (+); 4. (Z)−2-nonenal (+); 5. 5-Dimethyl-3 (2H)-isoxazolone (+)	Aromatic compounds catabolism
Gamagedara *et al.* 2012 [53]	*n*=12 (HC)	*n*=11	NS	NS	LC-MS/MS	NS/3	1. Taurine (+); 2. Phen (−), hippuric acid (−)	1. Cysteine sulfinic acid pathway; 2. Amino acid metabolism
Alberice *et al.* 2013 [54]	NA	*n*=48 NMIBC 26 LR 22 HR	≤60-70, >70	44 M 4 F	LC-MS CE-MS	684/27 75/27	1. His (+), Phe (+), Leu (+), Cys (−), Tyr (+), Trp (+); 2. Betaine (+); 3. Hypoxanthine (+), uric acid (−); 4. Sorbitol (−), galacticol (−), mannitol (−); 5. Carnitine derivatives (+); 6. N-acetyltryptophan (+); 7. Dopaquinone (−); 8. Carnosine (−); 9. Nε, Nε, Nε-trimethyllysine (−), Nε, Nε-dimethyllysine (−)	1, 8. Amino acid metabolism; 2. Phospholipid metabolism; 3. Purine metabolism; 4. Sugar alcohol metabolism; 5, 9. FA metabolism; 6. Tryptophan metabolism; 7. Tyrosine metabolism
Pasikanti *et al.* 2013 [55]	*n*=61 (HC)	*n*=38	68.3 ± 10.9	32 M 6 F	GCxGC-TOF-MS	533/43	1. Adipic acid (+); 2. Citrate (−); 3. Glycerol (−); 4. Melibiose (+); 5. Ribitol (−); 6. Gluconic acid (−) and derivatives (+); 7. Uridine (+), pseudouridine (+); 8. Anthranilic acid (+); 9. Coumaric acid (+); 10. Lactate (+); 11. Cyclopentane-1,2-diamine (+); 12. Dihydroxyacetone (−); 13. Erythritol (+); 14. Erythro-pentonic acid (+); 15. Ethyl tartrate (−); 16. Ethylmalonic acid (+); 17. Itaconic acid (+); 18. Heptadecanoic acid (+); 19.Hydroxybutyric acid (+); 20. Levulinic acid (−); 21. N-	1. Aromatic compounds catabolism; 2. TCA cycle; 3. TG metabolism; 4. Sugar metabolism; 5, 6. Pentose phosphate pathway; 7. Pyrimidine metabolism; 8, 21. Tryptophan metabolism; 9. Amino acid metabolism; 10. Glycolysis/Pyruvate metabolism; 11, 14-18, 20, 24-27, 29. NA; 12. Glycerolipid metabolism; 13. ABC transporter; 19. Ketone bodies metabolism; 22, 28. Tyrosine metabolism; 23. Biosynthesis of secondary metabolites

Table 2. Clinical metabolomic studies of BC and identified potential biomarkers for early diagnosis (Continued)

Sample type study [reference]	Study cohort size				Analytical platform	Total/Distinct metabolites found	Discriminatory metabolites	Dysregulated metabolic pathways
	Controls	BC	Age (years)	Gender				
							acetyl-anthranilic acid (+); 22. *p*-Cresol (+); 23. Pinene (−); 24. Ribonic acid (−); 25. Sebacic acid (−); 26. Sumiki's acid (−); 27. Talonic acid (+); 28. Vanillylmandelic acid (+), 3,4-Dihydroxyphenylpyruvate (+); 29. 2-Aminoisobutyric acid (+), 2-hydroxymalonic acid (+), 3-hydroxysebacic acid (+), 4-methoxycinnamic acid (+), 2,5-furandicarboxylic acid (−), 2-butanedioic acid (−), 2-dydroxyglutaric acid (−); 30. Unknown sugar (−)	
Jin *et al.* 2014 [56]	*n*=121 (52 w/ hematuria)	*n*=138 83 NMIBC 55 MIBC	65.6 ± 12.6	112 M 26 F	HPLC-QTOF-MS	NA/12	1. Acetyl-CoA (+), succinate (+), oxoglutarate (+); 2. Pyruvate (+), phosphoenolpyruvate (+); 3. Carnitine (+), isovalerylcarnitine (+), glutarylcarnitine (−), octenoylcarnitine (+), decanoylcarnitine (−); 4. Trimethyllysine (+); 5. Melatonin (−)	1. TCA cycle; 2. Glycolysis; 3. FA metabolism; 4. Amino acid metabolism; 5. Tryptophan metabolism
Wittmann *et al.* 2014 [57]	*n*=345 (78 w/hematuria, 178 w/o BC history, 89 HC)	*n*=95 14 LG 69 HG 12 NA	66.7 - 67.4	79 M 15 F 1 NS	UHPLC-MS/MS GC-MS	>1000/4	1. Palmitoyl sphingomyelin (+); 2. Lactate (+); 3. Adenosine (−); 4. Succinate (+)	1. Sphingolipid metabolism; 2. Glycolysis; 3. Purine metabolism; 4. TCA cycle
Shen *et al.* 2015 [58]	*n*=27 (21 HC, 6 post-surgery)	*n*=23	65.1 ± 13.3	34 M 16 F	UPLC-HRMS	>9000/6	1. Nicotinuric acid (+); 2. Trehalose (+); 3. Inosinic acid (−); 4. Ureidosuccinic acid (−); 5. Tetrapeptides AspAspGlyTrp (+) and GlyGlyAlaLys (−)	1. Nicotinate/nicotinamide metabolism, FA metabolism; 2. Sugar metabolism; 3. Purine metabolism; 4. Aspartate and Pyrimidine synthesis; 5. Amino acid metabolism

Note: (+) indicates increased metabolite concentration in BC, (−) indicates decreased metabolite concentration in BC.
[1]H NMR, proton nuclear magnetic resonance; Ala, alanine; Asn, asparagine; Asp, aspartate; BC, bladder cancer; CE, capillary electrophoresis; Cho, choline; ChoCC, choline-containing compounds; Cys, cystine; DMA, dimethylamine; F, female; FA, fatty acid; FLD, fluorescence detection; GC, gas-chromatography; Gln, glutamine; Glu, glutamate; Gly, glycine; GSH, glutathione (reduced form); HC, healthy controls; HG, high-grade bladder cancer; HILIC, hydrophilic-interaction liquid chromatography; His, histidine; HRMS, high resolution magic angle spinning; HPLC, high-performance liquid chromatography; HR, high risk recurrence; Ile, isoleucine; LC, liquid-chromatography; Leu, leucine; LG, low-grade bladder cancer; Lys, lysine; LR, low risk recurrence; M, male; MIBC, muscle invasive bladder cancer; MS, mass spectrometry; MTA, 5'-deoxy-5'-methylthioadenosine; NA, not available; NMIBC, nonmuscle invasive bladder cancer; NS, not specified; PAGN, phenylacetylglutamine; Phe, phenylalanine; SAM, *S*-adenosyl methionine; Ser, serine; TCA, tricarboxylic acid (cycle); TG, triglycerides; TOF, time-of-flight; TURBT, transurethral resection of the bladder tumor; Tyr, tyrosine; Trp, tryptophan; UDP, uridine diphosphate; UHPLC, ultra-high performance liquid chromatography; Val, valine; w/, with; w/o, without.

phenylalanine and histidine were found in BC tissues. Moreover, levels of S-adenosyl methionine (SAM) were also higher in BC tissues as well as aniline, a xenobiotic involved in bladder carcinogenesis,[69] whereas levels of taurine, and palmitic, lauric and oleic acids were decreased in BC compared with adjacent benign tissues.[44] The combination of these metabolites was able to distinguish both normal and benign tissues from BC tissues. Focusing on the same issue, Tripathi *et al.* (2013) performed a ^{1}H-HRMAS-NMR-based metabolomics analysis of benign and BC tissues, in which it was observed that BC tissues exhibited more metabolic abnormalities than benign ones, resulting in a clear differentiation between benign and BC patients. Nevertheless, this method was not able to distinguish various BC pathological stages (Ta/T1 versus $\geq$ T2), thus suggesting that metabolic features of different BC stages are difficult to separate. Twenty-two metabolites were found to be altered in BC tissues and cross-validation via targeted GC-MS analysis of the same tissue samples demonstrated the potential of those biomarkers for clinical diagnosis of BC. Potential metabolites included: triglycerides, which were decreased in BC tissues, and intermediates of glycolysis and TCA cycle, amino acids, taurine, GSH and choline-containing metabolites, which were elevated in BC tissues compared with benign tissues.[45]

Both studies substantiate the enhancement in amino acid metabolism, a consistent hallmark of cancer development, and in methylation, due to increased SAM levels,[44] consistent with increased DNA and protein synthesis. Another hallmark of cancer is the aberrant formation of cell membranes, pointed out by the altered levels of choline-containing compounds.[45] However, levels of taurine were contradictory between the studies, either found down-[44] or upregulated.[45]

Human biofluids studies

Blood plasma/serum studies. Two very promising ^{1}H-NMR-based metabolomics studies on serum were performed aiming at the distinction between BC patients and healthy controls, along with tumor stages. The one performed by Cao *et al.* (2012) examined the serum profile of patients with either high- or low-grade BC and included patients with urinary calculi (presenting with hematuria, as BC patients do) in the control group. Statistical analysis demonstrated that BC patients could be distinctly separated from both healthy controls and patients with calculi. Notably, metabolic profiles allowed for the distinction between low- and high-grade BC as well. It was observed that the serum levels of isoleucine/leucine, tyrosine, phenylalanine, choline, lactate, glycine and citrate were lower in BC patients compared with healthy subjects, whereas the levels of lipids and glucose were higher in BC patients. Compared with calculi patients, serum levels of metabolites in BC patients presented with the same trend of changes. Comparisons of the metabolite levels between low-grade and high-grade BC patients showed that levels of tyrosine, phenylalanine, lactate and glycine were higher in low-grade, whereas the levels of glucose were lower than those of high-grade BC.[47]

The second study was performed using serum samples from low- and high-grade BC patients, and healthy controls.[48] The statistical model was able to differentiate between BC patients and healthy individuals, as well as low- versus high-grade BC patients with high sensitivity and specificity. Six metabolites—dimethylamine (DMA), malonate, lactate, glutamine, histidine and valine—were found significantly different between the healthy and cancer groups (Table 2). Among those metabolites, glutamine was found at increased levels in low-grade BC, whereas DMA and malonate were found increased in high-grade BC. Of importance, in this study a new group of 106 suspected BC patients was also analyzed for external validation through a double-blind study, which confirmed the helpfulness of the metabolomic platform for the early diagnosis of BC.[48]

Serum-based metabolomics showed alterations in the energetic metabolism, particularly increased fatty acid synthesis and anaerobic glycolysis, as an alternative energy supply, and glycerophospholipids metabolism (higher levels of DMA), for cell membrane formation. These studies also evidenced the possibility to evaluate whether it is low- or high-grade BC. Lactate has been already associated with cancer progression.[43] Tyrosine, phenylalanine and glycine seem to play a major role in cancer pathogenesis,[70,71] glutamine metabolism provides additional energy for the uncontrolled proliferation[72] and malonate is assumed to inhibit the respiratory chain Complex II (succinate dehydrogenase) to help cancer cells against ROS production.[73]

Urine studies. A ^{1}H-NMR-based urine metabolomics study was carried out to investigate metabolic pathways perturbed in NMIBC.[50] Urines from BC patients, benign controls and healthy controls were analyzed (Table 2). A significant increase in the levels of taurine in BC samples compared with controls was observed, whereas levels of citrate, phenylalanine and hippuric acid were found to be decreased in BC. However, in this study, the stages of BC could not be determined based only on the indices of the altered metabolites and the sensitivity and specificity of this method were not validated. Based on this study, Gamagedara *et al.* (2012) developed a LC-MS/MS method to quantify taurine, phenylalanine and hippuric acid in urine, but including normalization to creatinine levels, being observed the same trend of changes. Taurine is a semiessential amino acid important as an osmolyte, antioxidant, free-radical scavenger, that protects cells against oxidative damage.[74,75] It also inactivates hypochlorous acid, a strong cytotoxic oxidant, by forming a stable complex, taurine-chloramine, which down-regulates immune responses leading to tumor development.[74] Taurine might be considered a good candidate for BC diagnosis since its increase in BC urine is concordant with its upregulation in BC tissues.[45] Nevertheless, higher concentrations of taurine have been reported in other cancers as well.[76–78]

Pasikanti *et al.* (2010) were also able to distinguish BC patients from healthy controls using another platform, GC-TOF-MS, with 100% sensitivity. Furthermore, and despite the need for a larger cohort to validate these results, the authors established a robust proof-of-principle for GC-TOF-MS-based metabolomics in staging and grading bladder tumours.[51] Among the 32 metabolites selected, 15 of them were significantly altered (Table 2), including valine, citrate, glycerol, melibiose, ribitol, gluconic acid, uridine, among others. A more recent study by the same group[55] is resumed in Table 2, and, interestingly, it suggested the potential role of kynurenine in the development of BC due to alterations in the tryptophan metabolism. Upregulation of tryptophan metabolic pathway leads to higher concentrations of anthranilic acid and *N*-acetyl-anthranilic acid,[55] kynurenine,[42,44] tryptophan[44,54] and *N*-acetyltryptophan.[54] These metabolites have been indicated as BC carcinogens, as their structure is similar to environmental ones.[79] Excessive tryptophan metabolites seem to play a role in suppressing antitumor immune responses, thus promoting cancer cells survival, through activation of aryl hydrocarbon receptor, which is involved in carcinogenesis.[80,81]

Two other studies carried out a LC-MS-based metabolomic analysis of urine[2,56] to build a screening model for BC diagnosis. Noteworthy, one of the studies included hematuria controls to exclude possible confounding effects of benign hematuria.[56] Various discriminative metabolites were identified in both studies (Table 2), particularly several altered carnitine species. Carnitine was found increased in BC urines, as well as isovalerylcarnitine, octenoylcarnitine and acetylcarnitine, consistent with the results from another BC urine study[54] and a BC tissue study.[44] Decreased levels of decanoylcarnitine, glutarylcarnitine and 2,6-dimethylhptanoyl carnitine were found in the cancer group. Additionally, the levels of carnitines were higher in MIBC than NMIBC,[56] which suggests that they may be correlated with BC aggressiveness, hence can be useful to distinguish grades of cancer. Furthermore, urinary levels of acetyl-CoA were also found elevated in BC patients.[56] Concerning carnitine, it is essential in the fatty acid β-oxidation, facilitating the entry of fatty acid into the mitochondria and acetyl-CoA is the end product of that oxidation. Therefore, excess acetyl-CoA and altered levels of carnitine species evidence that fatty acid β-oxidation is dysregulated in BC, consistent with other urine studies on this cancer.[54,57,58] Increased glycolysis and impaired TCA cycle might also lead to acetyl-CoA accumulation and its shift toward carnitine shuttle and fatty acid β-oxidation. Moreover, decreased citrate levels found in BC[42,47,50,51,55] reflect the increased conversion of citrate into fatty acids necessary for β-oxidation to support the rapid proliferation of cancer cells.

As expected, perturbations in purine and pyrimidine metabolic pathways were observed, due to enhanced cancer cells cycle activity. Pyrimidine metabolism was found upregulated due to elevated levels of uridine and pseudouridine in BC,[51,55] which are metabolites involved in RNA synthesis and reflect a greater energetic state of the tumor cells, and decreased levels of ureidosuccinic acid,[58] which may reflect its higher consumption for pyrimidine synthesis. Likewise, purine metabolism was found elevated with higher levels of hypoxanthine and decreased levels of uric acid,[54] adenosine and inosinic acid. In the normal breakdown of purine nucleotides, hypoxanthine is formed and converted to uric acid. In cancer cells, this pathway is dysregulated as purine synthesis is favored, therefore uric acid is diminished and hypoxanthine is accumulated for the *de novo* synthesis of purines. This bioprocess also explains the decreased levels of adenosine and inosinic acid in BC patients.

Conclusions and Future Perspectives

Overall, BC displays perturbations in several metabolic pathways, which involve essential biochemical reactions that generate energy, namely, glycolysis, TCA cycle, fatty acid β-oxidation, carnitine shuttle and amino acids metabolism, particularly the upregulation of tryptophan metabolic pathway.[34,42,44,48,54–56] This would be expected because cancer cells, in general, require more energy to grow and proliferate faster than normal cells. In turn, elevation of GSH levels was also noticed in this urological cancer,[40,45] consistent with its synthesis induction as a response to the higher oxidative stress in cancer cells. Concerning the alterations in glycerophospholipids metabolism, levels of choline and choline-containing compounds were found either elevated in BC tissues[45] or decreased in serum[47] and urine samples[2] from BC patients, which suggests either accumulation or excess consumption of these metabolites for the rapid cell membranes formation. Upregulation of purine and pyrimidine metabolism were also observed in urine,[51,54,55,57,58] tissues[44,45] and cell lines,[41] consistent with the higher rate of nucleic acid synthesis. The aforesaid studies indicate some promising metabolites that may be considered as good candidates for early detection of BC and disclose the main metabolic pathways dysregulated in BC, which are outlined in Figure 1.

Interestingly, some studies were able to distinguish between early and advanced stages of BC through the levels of certain metabolites.[43,47,48,54,56,58] The tendency of changes in some metabolites, including glucose, lactate,[43,47] tyrosine,[47,54] phenylalanine, glycine,[47] DMA, malonate, glutamine,[48] carnitine species,[54,56] hypoxanthine and uric acid,[54] was consistent with BC aggressiveness and stage of metastasis. These results suggest that different stages/grades of BC might generate distinct metabolic profiles, which might be due to the fact that cancer cells in advanced grades/stages require more energy for survival and continuous growing. Moreover, the replacement of the conventional two-pathway model of BC pathogenesis by complex molecular signatures that have been proposed based on large-scale genome-wide profiling studies (e.g., TCGA consortium, COSMIC) will require a shift in data analysis of the metabolomics profiling

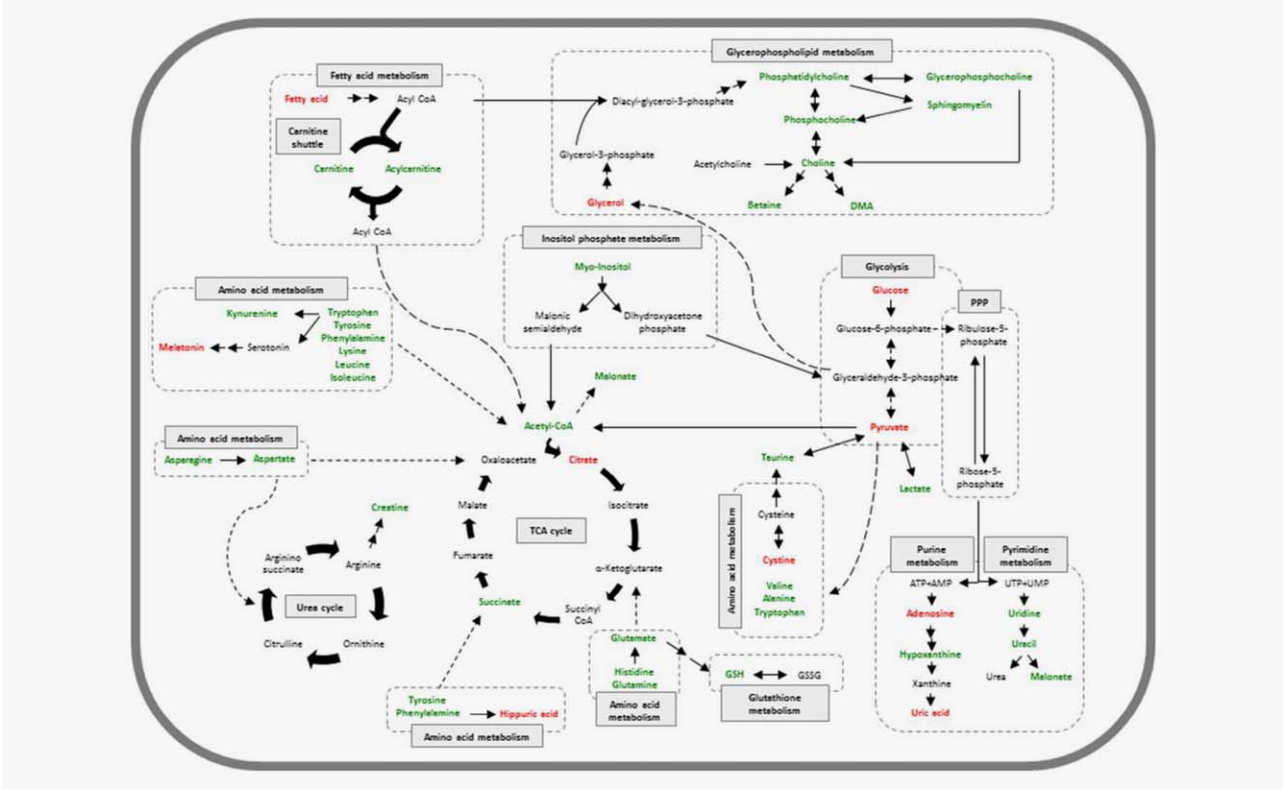

Figure 1. Schematic illustration of significant dysregulated metabolic pathways in BC, highlighting potential metabolite biomarkers. Metabolites downregulated are shown in red and those upregulated are shown in green. AMP, adenosine monophosphate; ATP, adenosine triphosphate; DMA, dimethylamine; GSH, glutathione (reduced form); GSSG, glutathione (oxidized form); PPP, pentose phosphate pathway; TCA, tricarboxylic acid (cycle); UMP, uridine monophosphate; UTP, uridine triphosphate. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

of BC. Indeed, future metabolomics studies will have to necessarily entail these new molecular signatures.[8]

Despite the progress of metabolomics in cancer research, there are still some challenges and limitations. For instance, the inadequate number of patients recruited in each study and small validation cohorts difficult the translation of metabolites to clinical practice. Additionally, only a few studies carried out a quantitative method to validate the key metabolites, because it can be challenging and time-consuming due to lack of standardized quantification methods.[34] Notably, there is some inconsistency and relatively little overlap between biomarkers identified among the studies, even when the same specimens were analyzed. This may be due to a series of factors, such as the work environmental conditions, different handling and storage of samples,[35] which may alter the already dynamic and sensitive metabolic profile, and the use of different analytical techniques that cover different metabolites. Another important challenge of metabolomic studies is that they are particularly affected by various confounding factors related to the patients, such as the genetic background, age, gender, lifestyle/diet, medication, surgical intervention and other pathological conditions.[15,34] This is particularly important in the case of urinary metabolomics since urine composition is highly affected by those factors and it may lead to false results.

In future metabolomic investigations, it is important to bear in mind that a good BC biomarker should be able not only to detect BC, but also to distinguish low- from high-grade BC, for a better diagnosis. Further developments and application of metabolomics will depend on various factors, such as the creation of a database containing the metabolites characteristic of BC, and their biochemical identities and interactions. It would be also interesting to integrate the metabolomic results with those obtained from other "omics" studies in order to cover all features of a certain cancer. Likewise, continuous development of high-throughput analytical platforms coupled with the improvement of bioinformatics tools will help overcome some limitations. Furthermore, standardization of the experimental design and analytical methods for biomarkers discovery is required, so that results become more comparable and validation of potential biomarkers can be achieved. Elimination of systematic effects of confounding variables on metabolite measurements needs further consideration as well, so that differences detected by multivariate analysis correspond to genuine changes in the levels of metabolites. Once these challenges and limitations are overcome, specific metabolites with potential clinical usefulness may be identified to aid in BC diagnosis, as well as in prognostication and long-term surveillance.

References

1. Burger M, Catto JW, Dalbagni G, *et al.* Epidemiology and risk factors of urothelial bladder cancer. *Eur Urol* 2013; 63:234–41.
2. Huang Z, Lin L, Gao Y, *et al.* Bladder cancer determination via two urinary metabolites: a biomarker pattern approach. *Mol Cell Prot* 2011;10: M111.007922.
3. Mitra AP, Cote RJ. Molecular pathogenesis and diagnostics of bladder cancer. *Annu Rev Pathol* 2009;4:251–85.
4. Pelucchi C, Bosetti C, Negri E, *et al.* Mechanisms of disease: the epidemiology of bladder cancer. *Nat Clin Pract Urol* 2006;3:327–40.
5. Kirkali Z, Chan T, Manoharan M, *et al.* Bladder cancer: epidemiology, staging and grading, and diagnosis. *Urology* 2005;66:4–34.
6. Eble JN, Stauter G, Epstein JI, *et al.* WHO Pathology and Genetics of Tumours of the Urinary System and Male Genital Organs. IARC Press: Lyon. 2004.
7. Kaufman DS, Shipley WU, Feldman AS. Bladder cancer. *Lancet* 2009;374:239–49.
8. Knowles MA, Hurst CD. Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity. *Nature Rev Cancer* 2015; 15:25–41.
9. Gazzaniga P, Gradilone A, de Berardinis E, *et al.* Prognostic value of circulating tumor cells in nonmuscle invasive bladder cancer: a CellSearch analysis. *Ann Oncol* 2012; 23:2352–6.
10. van der Poel HG, Debruyne FM. Can biological markers replace cystoscopy? An update. *Curr Opin Urol* 2001; 11:503–9.
11. Grossman HB, Soloway M, Messing E, *et al.* Surveillance for recurrent bladder cancer using a point-of-care proteomic assay. *JAMA* 2006; 295: 299–305.
12. Kim WJ, Park S, Kim YJ. Biomarkers in bladder cancer: present status and perspectives. *Biomarker Insights* 2007; 2:95–105.
13. Lokeshwar VB, Habuchi T, Grossman HB, *et al.* Bladder tumor markers beyond cytology: international consensus panel on bladder tumor markers. *Urology* 2005; 66:35–63.
14. Miremami J, Kyprianou N. The promise of novel molecular markers in bladder cancer. *Int J Mol Sci* 2014; 15:23897–908.
15. Urquidi V, Rosser CJ, Goodison S. Molecular diagnostic trends in urological cancer: biomarkers for non-invasive diagnosis. *Curr Med Chem* 2012; 19:3653–63.
16. Locasale JW, Cantley LC, Vander Heiden MG. Cancer's insatiable appetite. *Nat Biotechnol* 2009; 27:916–7.
17. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011; 144:646–74.
18. Aboud OA, Weiss RH. New opportunities from the cancer metabolome. *Clin Chem* 2013; 59:138–46.
19. Ramirez T, Daneshian M, Kamp H, *et al.* Metabolomics in toxicology and preclinical research. *Altex* 2013; 30:209–25.
20. Jobard E, Pontoizeau C, Blaise BJ, *et al.* A serum nuclear magnetic resonance-based metabolomic signature of advanced metastatic human breast cancer. *Cancer Lett* 2014; 343:33–41.
21. Nam H, Chung BC, Kim Y, *et al.* Combining tissue transcriptomics and urine metabolomics for breast cancer biomarker identification. *Bioinformatics* 2009; 25:3151–7.
22. Chen J, Zhou L, Zhang X, *et al.* Urinary hydrophilic and hydrophobic metabolic profiling based on liquid chromatography-mass spectrometry methods: differential metabolite discovery specific to ovarian cancer. *Electrophoresis* 2012; 33:3361–9.
23. Sreekumar A, Poisson LM, Rajendiran TM, *et al.* Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* 2009; 457:910–4.
24. Stenman K, Stattin P, Stenlund H, *et al.* H HRMAS NMR derived bio-markers related to tumor grade, tumor cell fraction, and cell proliferation in prostate tissue samples. *Biomarker Insights* 2011; 6:39–47.
25. Kim K, Taylor SL, Ganti S, *et al.* Urine metabolomic analysis identifies potential biomarkers and pathogenic pathways in kidney cancer. *Omics* 2011; 15:293–303.
26. Ganti S, Taylor SL, Kim K, *et al.* Urinary acylcarnitines are altered in human kidney cancer. *Int J Cancer* 2012; 130:2791–800.
27. Carrola J, Rocha CM, Barros AS, *et al.* Metabolic signatures of lung cancer in biofluids: NMR-based metabonomics of urine. *J Proteome Res* 2011; 10:221–30.
28. Erb G, Elbayed K, Piotto M, *et al.* Toward improved grading of malignancy in oligodendrogliomas using metabolomics. *Magn Reson Med* 2008; 59:959–65.
29. Srivastava NK, Pradhan S, Gowda GA, *et al.* In vitro, high-resolution 1H and 31P NMR based analysis of the lipid components in the tissue, serum, and CSF of the patients with primary brain tumors: one possible diagnostic view. *NMR Biomed* 2010; 23:22.
30. Chen F, Xue J, Zhou L, *et al.* Identification of serum biomarkers of hepatocarcinoma through liquid chromatography/mass spectrometry-based metabonomic method. *Anal Bioanal Chem* 2011; 401:1899–904.
31. Chen T, Xie G, Wang X, *et al.* Serum and urine metabolite profiling reveals potential biomarkers of human hepatocellular carcinoma. *Mcp* 2011; 10:M110.004945.
32. Spratlin JL, Serkova NJ, Eckhardt SG. Clinical applications of metabolomics in oncology: a review. *Clin Cancer Res* 2009; 15:431–40.
33. Yoon M, Campbell JL, Andersen ME, *et al.* Quantitative in vitro to in vivo extrapolation of cell-based toxicity assay results. *Crit Rev Toxicol* 2012; 42:633–52.
34. Chan EC, Pasikanti KK, Hong Y, *et al.* Metabonomic profiling of bladder cancer. *J Proteome Res* 2015; 14:587–602.
35. Monteiro MS, Carvalho M, Bastos ML, *et al.* Metabolomics analysis for biomarker discovery: advances and challenges. *Curr Med Chem* 2013; 20:257–71.
36. Zhang J, Wei S, Liu L, *et al.* NMR-based metabolomics study of canine bladder cancer. *Biochim Biophys Acta* 2012; 1822:1807–14.
37. Gika HG, Theodoridis GA, Wilson ID. Liquid chromatography and ultra-performance liquid chromatography-mass spectrometry fingerprinting of human urine: sample stability under different handling and storage conditions for metabonomics studies. *J Chromatogr* 2008; 1189:314–22.
38. Shulaev V. Metabolomics technology and bioinformatics. *Brief Bioinf* 2006; 7:128–39.
39. Nagrath D, Caneba C, Karedath T, *et al.* Metabolomics for mitochondrial and cancer studies. *Biochim Biophys Acta* 2011; 1807:650–63.
40. Pendyala L, Velagapudi S, Toth K, *et al.* Translational studies of glutathione in bladder cancer cell lines and human specimens. *Clin Cancer Res* 1997; 3:793–8.
41. Pasikanti KK, Norasmara J, Cai S, *et al.* Metabolic footprinting of tumorigenic and nontumorigenic uroepithelial cells using two-dimensional gas chromatography time-of-flight mass spectrometry. *Anal Bioanal Chem* 2010; 398:1285–93.
42. Dettmer K, Vogl FC, Ritter AP, *et al.* Distinct metabolic differences between various human cancer and primary cells. *Electrophoresis* 2013; 34: 2836–47.
43. Conde VR, Oliveira PF, Nunes AR, *et al.* The progression from a lower to a higher invasive stage of bladder cancer is associated with severe alterations in glucose and pyruvate metabolism. *Exp Cell Res* 2015; 335:91–8.
44. Putluri N, Shojaie A, Vasu VT, *et al.* Metabolomic profiling reveals potential markers and bioprocesses altered in bladder cancer progression. *Cancer Res* 2011; 71:7376–86.
45. Tripathi P, Somashekar BS, Ponnusamy M, *et al.* HR-MAS NMR tissue metabolomic signatures cross-validated by mass spectrometry distinguish bladder cancer from benign disease. *J Proteome Res* 2013; 12:3519–28.
46. Lin L, Huang Z, Gao Y, *et al.* LC-MS-based serum metabolic profiling for genitourinary cancer classification and cancer type-specific biomarker discovery. *Proteomics* 2012; 12:2238–46.
47. Cao M, Zhao L, Chen H, *et al.* NMR-based metabolomic analysis of human bladder cancer. *Anal Sci* 2012; 28:451–6.
48. Bansal N, Gupta A, Mitash N, *et al.* Low- and high-grade bladder cancer determination via human serum-based metabolomics approach. *J Proteome Res* 2013; 12:5839–50.
49. Issaq HJ, Nativ O, Waybright T, *et al.* Detection of bladder cancer in human urine by metabolomic profiling using high performance liquid chromatography/mass spectrometry. *J Urol* 2008; 179: 2422–6.
50. Srivastava S, Roy R, Singh S, *et al.* Taurine - a possible fingerprint biomarker in non-muscle invasive bladder cancer: a pilot study by 1H NMR spectroscopy. *Cancer Biomarker* 2010; 6: 11–20.
51. Pasikanti KK, Esuvaranathan K, Ho PC, *et al.* Noninvasive urinary metabonomic diagnosis of human bladder cancer. *J Proteome Res* 2010; 9: 2988–95.
52. Jobu K, Sun C, Yoshioka S, *et al.* Metabolomics study on the biochemical profiles of odor elements in urine of human with bladder cancer. *Biol Pharm Bull* 2012; 35:639–42.
53. Gamagedara S, Shi H, Ma Y. Quantitative determination of taurine and related biomarkers in urine by liquid chromatography-tandem mass spectrometry. *Anal Bioanal Chem* 2012; 402:763–70.
54. Alberice JV, Amaral AF, Armitage EG, *et al.* Searching for urine biomarkers of bladder cancer recurrence using a liquid chromatography-mass spectrometry and capillary electrophoresis-mass spectrometry metabolomics approach. *J Chromatogr* 2013; 1318:163–70.

55. Pasikanti KK, Esuvaranathan K, Hong Y, *et al.* Urinary metabotyping of bladder cancer using two-dimensional gas chromatography time-of-flight mass spectrometry. *J Proteome Res* 2013; 12:3865–73.
56. Jin X, Yun SJ, Jeong P, *et al.* Diagnosis of bladder cancer and prediction of survival by urinary metabolomics. *Oncotarget* 2014; 5:1635–45.
57. Wittmann BM, Stirdivant SM, Mitchell MW, *et al.* Bladder cancer biomarker discovery using global metabolomic profiling of urine. *PloS One* 2014; 9:e115870.
58. Shen C, Sun Z, Chen D, *et al.* Developing urinary metabolomic signatures as early bladder cancer diagnostic markers. *Omics* 2015; 19:1–11.
59. Pelicano H, Carney D, Huang P. ROS stress in cancer cells and therapeutic implications. *Drug Resist Updates* 2004; 7:97–110.
60. Giralt M, Lafuente A, Pujol F, *et al.* Enhanced glutathione S-transferase activity and glutathione content in human bladder cancer. Followup study: influence of smoking. *J Urol* 1993; 149: 1452–4.
61. Bayet-Robert M, Loiseau D, Rio P, *et al.* Quantitative two-dimensional HRMAS 1H-NMR spectroscopy-based metabolite profiling of human cancer cell lines and response to chemotherapy. *Magn Reson Med* 2010; 63: 1172–83.
62. Hines KM, Ballard BR, Marshall DR, *et al.* Structural mass spectrometry of tissue extracts to distinguish cancerous and non-cancerous breast diseases. *Mol Biosyst* 2014; 10:2827–2837.
63. Ganti S, Taylor SL, Abu Aboud O, *et al.* Kidney tumor biomarkers revealed by simultaneous multiple matrix metabolomics analysis. *Cancer Res* 2012; 72:3471–9.
64. Huang Q, Tan Y, Yin P, *et al.* Metabolic characterization of hepatocellular carcinoma using nontargeted tissue metabolomics. *Cancer Res* 2013; 73:4992–5002.
65. Imperiale A, Elbayed K, Moussallieh FM, *et al.* Metabolomic pattern of childhood neuroblastoma obtained by (1)H-high-resolution magic angle spinning (HRMAS) NMR spectroscopy. *Pediatr Blood Cancer* 2011; 56:24–34.
66. Monleon D, Morales JM, Gonzalez-Darder J, *et al.* Benign and atypical meningioma metabolic signatures by high-resolution magic-angle spinning molecular profiling. *J Proteome Res* 2008; 7: 2882–8.
67. Rocha CM, Barros AS, Goodfellow BJ, *et al.* NMR metabolomics of human lung tumours reveals distinct metabolic signatures for adenocarcinoma and squamous cell carcinoma. *Carcinogenesis* 2015; 36:68–75.
68. Feron O. Pyruvate into lactate and back: from the Warburg effect to symbiotic energy fuel exchange in cancer cells. *Radiother Oncol* 2009; 92:329–33.
69. Carreon T, Hein MJ, Viet SM, *et al.* Increased bladder cancer risk among workers exposed to o-toluidine and aniline: a reanalysis. *Occup Environ Med* 2010; 67:348–50.
70. Letellier S, Garnier JP, Spy J, *et al.* Determination of the L-DOPA/L-tyrosine ratio in human plasma by high-performance liquid chromatography. Usefulness as a marker in metastatic malignant melanoma. *J Chromatogr* 1997; 696:9–17.
71. Duarte IF, Rocha CM, Barros AS, *et al.* Can nuclear magnetic resonance (NMR) spectroscopy reveal different metabolic signatures for lung tumours? *Virchows Arch* 2010; 457:715–25.
72. Dang CV. Glutaminolysis: supplying carbon or nitrogen or both for cancer cells? *Cell Cycle* 2010; 9:3884–6.
73. Fernandez-Gomez FJ, Galindo MF, Gomez-Lazaro M, *et al.* Malonate induces cell death via mitochondrial potential collapse and delayed swelling through an ROS-dependent pathway. *Br J Pharmacol* 2005; 144:528–37.
74. Rosado JO, Salvador M, Bonatto D. Importance of the trans-sulfuration pathway in cancer prevention and promotion. *Mol Cell Biochem* 2007; 301:1–12.
75. Redmond HP, Stapleton PP, Neary P, *et al.* Immunonutrition: the role of taurine. *Nutrition* 1998; 14:599–604.
76. Bezabeh T, Odlum O, Nason R, *et al.* Prediction of treatment response in head and neck cancer by magnetic resonance spectroscopy. *AJNR* 2005; 26:2108–13.
77. Beckonert O, Monnerjahn J, Bonk U, *et al.* Visualizing metabolic changes in breast-cancer tissue using 1H-NMR spectroscopy and self-organizing maps. *NMR Biomed* 2003; 16:1–11.
78. De Micheli E, Alfieri A, Pinna G, *et al.* Extracellular levels of taurine in tumoral, peritumoral and normal brain tissue in patients with malignant glioma: an intraoperative microdialysis study. *Adv Exp Med Biol* 2000; 483:621–5.
79. Sidransky H. Tryptophan and carcinogenesis: review and update on how tryptophan may act. *Nutr Cancer* 1997; 29:181–94.
80. Chung KT, Gadupudi GS. Possible roles of excess tryptophan metabolites in cancer. *Environ Mol Mutagen* 2011; 52:81–104.
81. Opitz CA, Litzenburger UM, Sahm F, *et al.* An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. *Nature* 2011; 478:197–203.

3. AIMS OF THE EXPERIMENTAL WORK

Concerning BC, there have been great attempts to improve diagnostic strategies and several potential markers have been described over the years or are under investigation, and some of them have already been validated and approved to be used in clinical practice (58, 123-125). Nevertheless, most still fail the early detection of this malignancy, presenting no advantages over classical diagnosis methods (126), and their high cost limit their application in clinical practice (123). Therefore, the need to find more accurate, disease-specific and easily detectable and economical biomarkers is urgent.

BC metabolomics studies, as described in the previous chapter, have shown promising results and good potential biomarkers, but there is still a lot to be done. With the hope of contributing to the search of BC biomarkers and to the better understanding of BC metabolism, the present work focused on performing an untargeted metabolomic study, using *in vitro* BC cell lines. Despite its aforementioned disadvantages, the *in vitro* model system was chosen due to the lower complexity in processing the samples and in controlling the experimental settings, compared with other matrices, and because it provides a first insight on the metabolic alterations that may occur in the organism in the presence of BC.

The overarching goals of the present work are as follow:

a) Application of HS-SPME/GC-MS based metabolomics for the metabolic profiling of 3 different tumorigenic bladder cell lines and of a non-tumorigenic cell line to search for significant differences among them.
b) Evaluation of two different pHs (pH 2 and 7) in order to infer which pH is optimum for VOCs extraction that can be applied in future *in vitro* metabolomic studies.
c) Screening, statistical analysis and possible identification of VOCs that contribute to the separation of the cell lines.
d) Evaluation of the potential of the metabolic signature found for discrimination of different types and grades of BC.

VOCs were analysed due to their great potential as biomarkers and the relatively easy translatability to other biological matrices (such as urine or exhaled breath). In addition, analysis of VOCs using HS-SPME/GC-MS requires simpler and faster sample preparation, which can be seen as an advantage in a clinical environment. Furthermore, to our knowledge, there is no BC metabolomic study that performed the VOCs profiling in *in vitro* cell lines, making this work pivotal in that matter.

4. EXPERIMENTAL WORK – MATERIAL AND METHODS

4.1. Chemicals

Minimum Essential Medium Eagle (MEM) and Ham's F-12 K nutrient mixture powder, both supplemented with L-glutamine, were purchased from Sigma-Aldrich (St. Louis, MO, USA). Penicillin, streptomycin and trypsin were purchased from Invitrogen (Karlsruhe, Germany); fetal bovine serum (FBS) was purchased from PAA laboratories GmbH (Colbe, Germany); DMSO (99.0%) and hydrochloric acid (HCl) were obtained from Merck (Darmstadt, Germany). Sodium chloride (NaCl, ≥99.5%), sodium hydroxide (NaOH, ≥97%), Folin-Ciocalteu's phenol reagent, bovine serum albumin (BSA) and 4-fluorobenzaldehyde (98%) were purchased from Sigma-Aldrich (St. Louis, MO, USA), and methanol (99.9%) from VWR (Leuven, Belgium).

All chemicals were of analytical grade and were dissolved in deionized water unless otherwise indicated.

4.2. Cell lines and culture conditions

The BC cell lines, namely J82, Scaber and 5637 were kindly provided by Instituto Português de Oncologia do Porto (IPO). The non-tumorigenic SV-HUC 1 was obtained from American Type Culture Collection (ATCC; Manassas, VA, USA). General characteristics of all cell lines are summarized in Table 1. All cancer cell lines were cultured in plastic T75 culture flasks containing 15 mL of MEM whereas SV-HUC-1 was cultured in plastic T75 culture flasks containing 15 mL of Ham's F-12 K medium. Both medium were prepared as indicated by the manufacturer and supplemented with 10% FBS and 50 units.mL^{-1} penicillin/50 µg.mL^{-1} streptomycin. To minimize possible effects of the used medium on metabolic footprints, the culture conditions were kept constant throughout the entire study. All cell lines were incubated at 37ºC and 5% CO_2.

Table 1. Characteristics of BC and normal cell lines used in this study.

	SV-HUC-1	5637	J82	Scaber
Organism	*Homo sapiens*	*Homo sapiens*	*Homo sapiens*	*Homo sapiens*
Gender/Age (years)	male/11	male/68	male/58	male/58
Tissue	ureter, uroepithelium	urinary bladder	urinary bladder	urinary bladder
Cell type	epithelial SV40 immortalized	primary tumour	primary tumour	primary tumour
Morphology	epithelial	epithelial	epithelial	epithelial
Culture properties	adherent	adherent	adherent	adherent
Tumorigenic	no	yes	yes	yes
Disease	-	TCC	TCC	SCC
Grade	-	II	III/IV	III/IV
Stage	-	NA	pT3	pT4

NA, not available; SCC; squamous cell carcinoma; TCC, transitional cell carcinoma

4.3. Collection of VOCs from extracellular medium

The experiments were carried out during 5 passages (passages 4 to 8 in the case of cancer cell lines and passages 6 to 10 in the case of SV-HUC-1) and in triplicate for each passage, resulting in a total number of 60 experiments. After cells' growth to maximum confluence was achieved, medium from each T75 culture flask was discarded and 15 mL of fresh medium was added and incubated at 37 ºC for 48 h. Afterwards, the extracellular medium was transferred to falcons on ice and centrifuged for 4 minutes at 2000 x*g* at 4 ºC. The supernatant was divided into 2 aliquots of 7 mL, for further samples analysis at two different pHs, and stored at -80 ºC until analysis. The remaining pellets were used for posterior protein quantification (described in 4.4).

T75 culture flasks containing only medium (controls) were treated in an identical manner. Controls medium was submitted to the same collection and storage conditions.

4.4. Protein quantification

Pellets were used to quantify protein content so that cellular density in each flask could be verified. It was applied the Lowry method (127), which is based in two reactions: (1) production of Cu^{+} due to the reaction between peptide bonds of proteins with cupper under alkaline conditions; (2) reaction of Cu^{+} with the Folin-Ciocalteau reagent that leads

to its reduction, and, consequently, it results in a strong blue colour solution that absorbs between 650 and 750 nm (128). The method's sensitivity is around 0.1 mg of protein.mL^{-1}.

Firstly, a calibration curve was obtained using different concentrations of a BSA stock solution 2000 $\mu g.mL^{-1}$, prepared in NaOH 1M. Secondly, different dilutions of the pellets in NaOH 1M were tested, so that the dilution with the most reproducible results and within the calibration curve was chosen. Afterwards, 50 µL of BSA, 100 µL of Solution A (Na_2CO_3 2%, $CuSO_45H_2O$ 1%, $KNaC_4H_4O_6$ 2%) and each pellet sample were added in each well of a 96-well plate. The plate was then covered for 10 min, after which 100 µL of Solution B (Folin-Ciocalteau reagent 1M) was added and the plate covered again for 20 min. Absorbances were measured at 750 nm in a Power Wave X microplate reader (Bio-Tek Instruments, Inc.; Winooski, VT, USA).

4.5. Sample preparation for HS-SPME/GC-MS analysis

4.5.1. Metabolic profiling of VOCs

Stored samples were thawed slowly in ice to minimize the loss of volatile compounds. Two mL of each sample were collected to a 10 mL glass vial, capped with a polytetrafluoroethylene (PTFE) septum and a screw cap, containing 0.59 g NaCl and 10 µL of internal standard (IS) 4-flurobenzaldehyde (Millipore, Bedford, MA, USA) with final concentration of 10 $\mu g.mL^{-1}$. As aforementioned, medium of each cell line and controls were analyzed at two different pHs (pH 2 and 7) in order to evaluate which pH is optimum for VOCs extraction. For pH 7 samples, there was no need to adjust the pH. After measuring all samples pH, the median pH was 7.252 ± 0.088. In the case of pH 2 samples, 70 - 75 µL of a 5 M HCl solution was added to adjust the pH, whose median was 2.076 ± 0.045.

4.5.2. QC samples

QCs were prepared as a pool of all samples in the study, where 200 µL of each sample was gathered in one single sample. Subsequently, the sample was divided into aliquots to avoid the constant freezing and thawing.

4.6. HS-SPME/GC-MS analysis: equipment and conditions

A general illustration of HS-SPME method coupled with GC is represented in figure 2. The HS-SPME measures were performed using a Combi-PAL autosampler (Varian Pal

Autosampler, Switzerland) and the Cycle Composer software (CTC Analytics System Software, Switzerland). Headspace (HS) volatiles were extracted by exposing a 65 µm film thickness polydimethylsiloxane/divinylbenzene (PDMS/DVB) fiber (Supelco Inc., Bellefonte, PA), previously conditioned at 250 °C for 30 minutes. The HS-SPME method was developed and optimized previously by our group (118). The HS was generated after 5 min. of incubation and 20 min. of extraction, at 45 °C.

The chromatographic analysis of the VOCs extracted from the extracellular medium were performed on a SCION^{TM} SQ (single quadrupole) 436-GC-MS system (Bruker Daltonics, Fremont, CA) equipped with a SCION SQ ion trap mass detector and a Bruker Daltonics MS workstation software (version 8.2). A capillary column VF-5 ms (30 m x 0.25 mm x 0.25 µm) from Varian was used for the chromatographic separation. Helium C-60 (Gasin, Portugal) was used as the carrier gas at a constant flow rate of 1.0 mL/min. The injection was in splitless mode and the injector temperature was 230 °C (held for 20 min.). As for the oven temperature, it was held for 1 min at 40 °C, followed by an increase at a rate of 5 °C/min until reaching 250 °C (held for 5 min) and finally an increase of 5 °C/min to 300 °C (held for 1 min). The MS detector was operated in EI mode. Data acquisition was performed in full scan mode with a mass range between 40-350 m/z at a scan rate of 6 scan/s.

QCs were analysed under the same conditions, on every 15 samples.

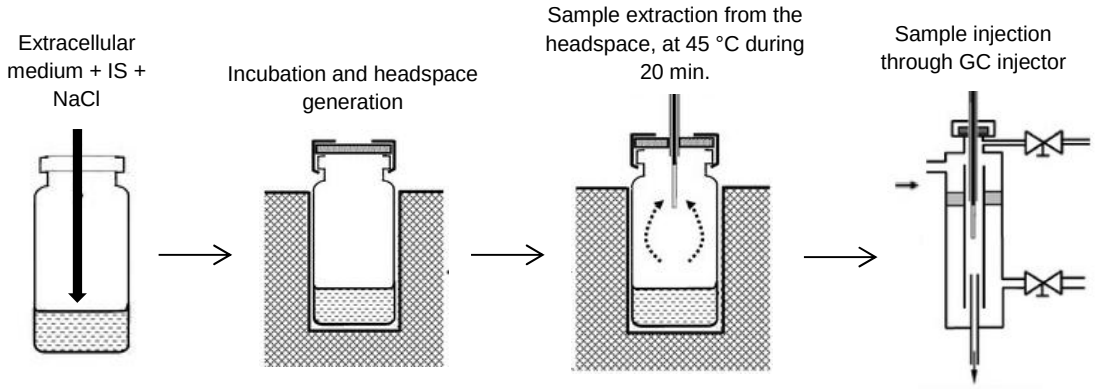

Figure 2. Schematic representation of a typical HS-SPME performed in the present work. Illustration was based on figure 3 from Pragst *et. al* (129).

4.7. GC-MS data pre-processing

VOCs identification in HS-SPME/GC-SQ/MS chromatograms of extracellular medium was performed based on the comparison of their retention times (RTs), Kovats retention index (RI) and mass spectra from the NIST mass spectral library (2014). The RI

of the analytes was calculated according to the retention times obtained for a solution of *n*-alkanes (C8-C20) series, which was analysed under the same chromatographic conditions as the set of samples. Only a reverse match of 700 or above was considered in the compound identification whereas compounds for which no satisfactory match was found were listed as "Unknown i" (i=1,2,3…) according to their crescent RTs. The HMDB (130) and KEGG (131) databases were also consulted to aid in the identification of compounds and interpretation of possible disturbed metabolic pathways.

All raw data files obtained from HS-SPME/GC-SQ/MS were exported as CDF files and pre-processed in order to convert instrumental data sets into a manageable format for data analysis and remove any bias such as background, noise and RT fluctuations over a set of samples. Data pre-processing was performed using the software MZmine 2.21 (132) and consisted in baseline correction, peak detection, chromatogram deconvolution, alignment and normalization. Baseline correction was applied to remove random noise and baseline shift in the raw data. Peak detection enabled the identification of all individual peaks caused by true ions avoiding detection of false positives. Chromatogram deconvolution was used to separate two or more co-eluting peaks in GC chromatogram using MS spectra. Alignment consisted in matching equivalent peaks across multiple samples. The parameters used in these steps were set as follow: RT range 2.8-34.0 min; m/z range 50-250; MS data noise level 1.0×10^4; m/z tolerance 0.5 or 5 ppm; chromatogram baseline level 1.0×10^3; peak duration range 0.02-0.30 min. After pre-processing steps, artefact peaks such as GC contaminants (e.g., cyclosiloxanes, siloxanes and phthalates) were removed from data matrix, as well as all peaks with relative standard deviation (RSD) $\leq$ 30 % across all QCs, as an indicator of poor reproducibility and repeatability of the analysis. Data was subsequently normalized by total area (TA) by dividing each peak area in the chromatogram by the total peak areas of the respective chromatogram. TA normalization is useful for correcting minor variations that derive from sample preparation and analysis (37). The resulting m/z-RT pairs, sample names and normalized peak areas were subjected to statistical analysis.

4.8. Statistical analysis

The final matrix was imported to SIMCA-P 13.0.3 (Umetrics, Umea, Sweden) and scaled to pareto (Par), which applies the square root of the standard deviation as a scaling factor to reduce the relative importance of small variables that might be from irrelevant peaks (133). PCA was first applied to detect trends and outliers, followed by

PLS-DA, used to discriminate between classes and identify the metabolic signature associated with a specific sample class. Model robustness was initially evaluated in terms of R^2X (variance explained by the **X** matrix, i.e. GC-MS data), R^2Y (variance explained by the **Y** matrix, i.e., sample class) and Q^2 (goodness of prediction or prediction power), a set of parameters obtained by 7-fold cross validation in SIMCA-P 13.0.3 software. The most discriminative variables (m/z-RT pairs) responsible for group separation were assessed through inspection of PLS-DA loadings scatter plots and the corresponding variable importance to the projection (VIP) of each variable. Only variables with VIP>1 were considered as important for group discrimination. Monte Carlo cross validation (MCCV) (7 blocks, 500 runs) was also carried out to evaluate the robustness of the PLS-DA models using a software developed in the University of Aveiro (134). The prediction power (Q^2) and confusion matrices of each model were computed for original and permuted (randomly splitting sample group) models. PLS-DA models were considered robust when minimal overlapping of the distribution of original and permuted Q^2 was obtained (135, 136). Classification rates, specificity (spec.) and sensitivity (sens.) were recovered for each PLS-DA model through a receiver operating characteristic (ROC) map.

The statistical significance of relevant compounds identified in the loadings scatter plots was computed using GraphPad Prism version 6 (GraphPad Software, San Diego, CA, USA). Firstly, Shapiro-Wilk normality test was applied to determine if data assumed a Gaussian distribution. For normally distributed data, an unpaired Student's t-test with Welch correction was applied, whereas for non-normally distributed data, an unpaired Mann-Whitney test was used. For each model, the discriminative compounds were considered statistically significant when *p*-value < 0.05 (confidence level 95%). Bonferroni correction (137) was used to adjust *p*-values for multiple comparisons by setting the significance cut-off to α/n, where $\alpha = 0.05$ and n represents the number of compounds simultaneously tested in univariate statistical analysis. Furthermore, for each significant compound, the effect size (calculated as described in Berben *et. al* (138)), percentage of variation and uncertainty were determined. Finally, PCA and PLS-DA were performed using the set of statistically significant compounds in order to further confirm the robustness of the metabolic signature found for each cancer cell type compared with the normal cell line. In order to confirm the robustness of the PLS-DA including only the discriminative metabolites, MCCV was also performed. Moreover, an example of a correlation network analysis (Gephi) of VOCs from both pHs selected as significant in discriminating the cancer cell line J82 (HG BC) from the normal cell line SV-HUC-1, based on Spearman's correlation coefficients ($|r| \geq 0.8$ and $p < 0.01$) was carried out in order to better understand how compounds would correlate.

5. Results

5.1. Data pre-processing and chemometric analysis

As aforementioned, in any *in vitro* metabolomic study it is necessary to take into account the cellular density, because it will affect the metabolite concentration of the samples. Although maximum cell confluence was assured before sample collection, protein quantification was measured to confirm if cells density was identical in all culture flaks. Indeed, data obtained for protein quantification was identical among all bladder cell lines (Supplementary Table S1, Appendix), showing that there was little difference in cell density per culture flask. Hence, data was not normalized for protein concentration but rather by chromatograms TA, as described previously.

Two different pHs were analysed to evaluate which one would yield better results and, therefore, be most appropriate in future *in vitro* metabolomic studies for VOCs extraction. In general, a variety of chemical compounds was detected in both pH 2 and pH 7, such as alcohols, aldehydes, ketones, alkanes, esters, carboxylic acids and monoterpenoids.

5.1.1. GC-MS analysis of the samples at pH 7

The GC-MS analysis of the extracellular medium at pH 7 consistently resulted in a total of 191 peaks. Representative full scan chromatograms of each cell line are presented with some of the most abundant compounds, as well as their respective controls medium (figure 3).

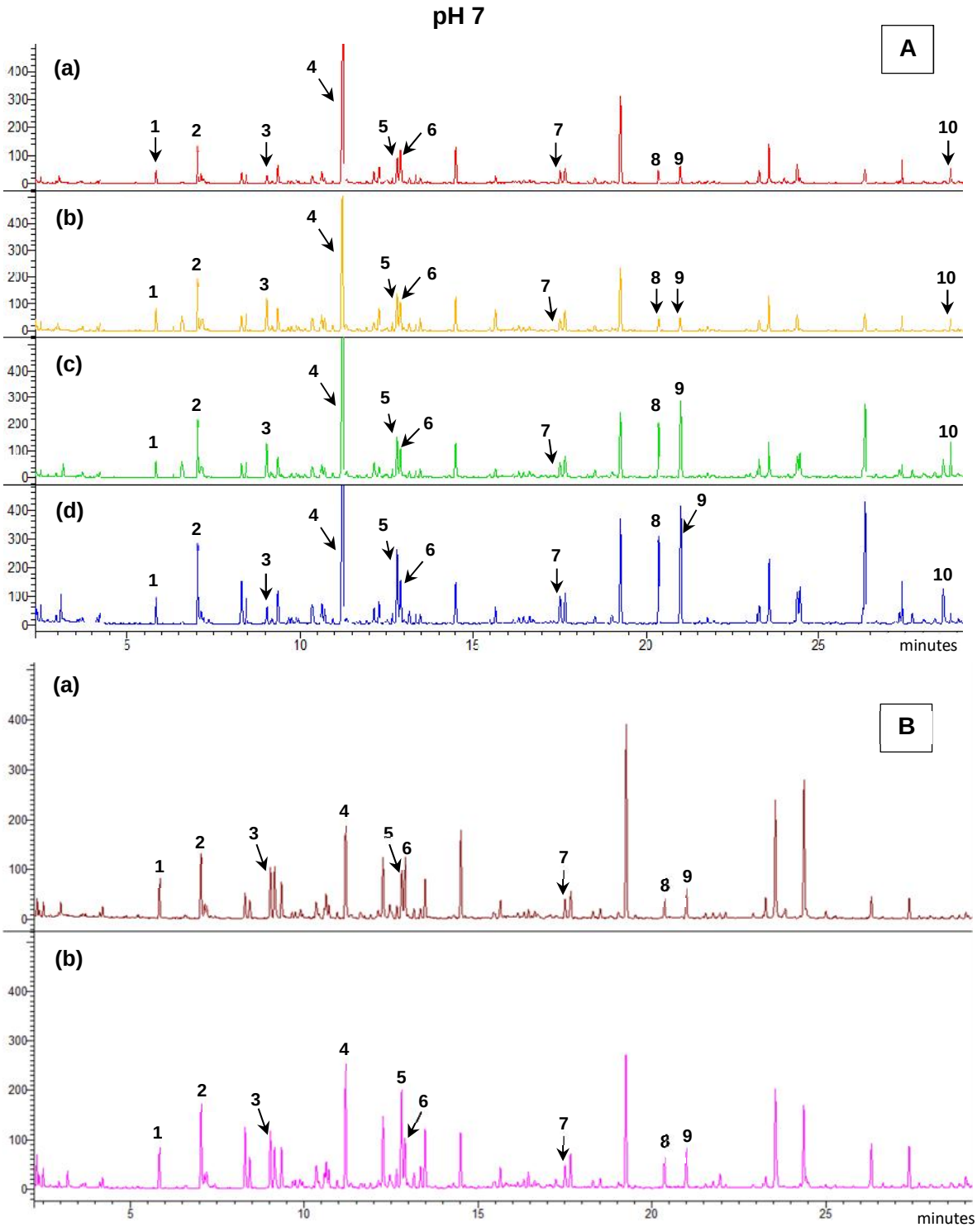

Figure 3. A. Representative full scan chromatograms obtained from the analysis of the extracellular medium of (a) SV-HUC-1, (b) 5637, (c) J82 and (d) Scaber, at pH 7. B. Representative full scan chromatograms obtained from the analysis of the controls medium of (a) SV-HUC-1 and (b) 5637, J82 and Scaber, at pH 7. Legend: 1. 2-4-dimethylhept-1-ene; 2. cyclohexanol; 3. 4-fluorobenzaldehyde (IS); 4. 2-ethyl-1-hexanol; 5. 4-methylbenzaldehyde; (6) 2-phenylpropan-2-ol; 7. 1-phenoxypropan-2-ol; 8. unknown; 9. 2-ethyl-3-hydroxyhexyl 2-methylpropanoate; 10. 2-pentadecanone. Note: the compounds are characterized by their IUPAC names.

With the aim of confirming that analytical variation was minimal and the method was reproducible, a pool of samples was analysed (QCs). Indeed, all QCs were closely clustered and centred in the PCA-X score scatter plot (figure 4), demonstrating the reproducibility of the method. Since the analytical variation was not significant, it is possible to infer that variations in the metabolites concentrations were mainly due to differences between the bladder cell lines.

In figure 4 it is also possible to observe a clear separation on the first principal component (PC1) between the culture medium of the cell lines and the controls.

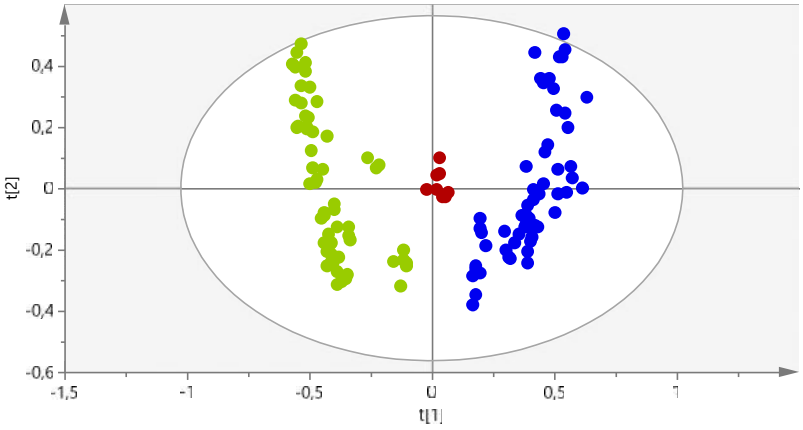

Figure 4. PCA-X score scatter plot obtained for the HS-SPME/GC-MS chromatograms of all samples, at pH 7, namely the medium of the controls (n=60, ●), QCs (n=8, ●) and the extracellular medium of all cell lines (n=60, ●). The ellipses indicate the 95% confidence limit of the model.

PCA-X and PLS-DA scores scatter plots were performed considering only the extracellular medium of the bladder cell lines, in order to study which compounds were responsible for the separation. The multivariate statistical models showed a clear separation between the extracellular medium of cancer cell lines and the one from the normal cell line, on the first principal component (PC1) (figures 5a to 5f). Values for R^2X, R^2Y and Q^2 for each PLS-DA model are also presented in the figures. Of note, one sample from SV-HUC-1 group was excluded from the PLS-DA model and from the following statistical analysis since it was considered an outlier, a decision based on the facts that the intensity of its chromatographic peaks was divergent from the rest of the samples and, when included in the PLS-DA models, it significantly diminished the Q^2 values. Moreover, in figures 6a and 6b, it is also illustrated an example of a PLS-DA model (Scaber vs SV-HUC-1) and its respective loading scatter plot that enabled the identification of the metabolites responsible for group separation (VIP>1), and their changes were verified by visual interpretation of boxplots performed on Graphpad and by

determining the % of variation and effect size values. This method was performed for all cell lines comparisons. It is also represented the Q^2 distribution and ROC plot of true and permuted classes in figures 6c and 6d, respectively, obtained by MCCV (Table 2) for the PLS-DA model exemplified in figure 6a.

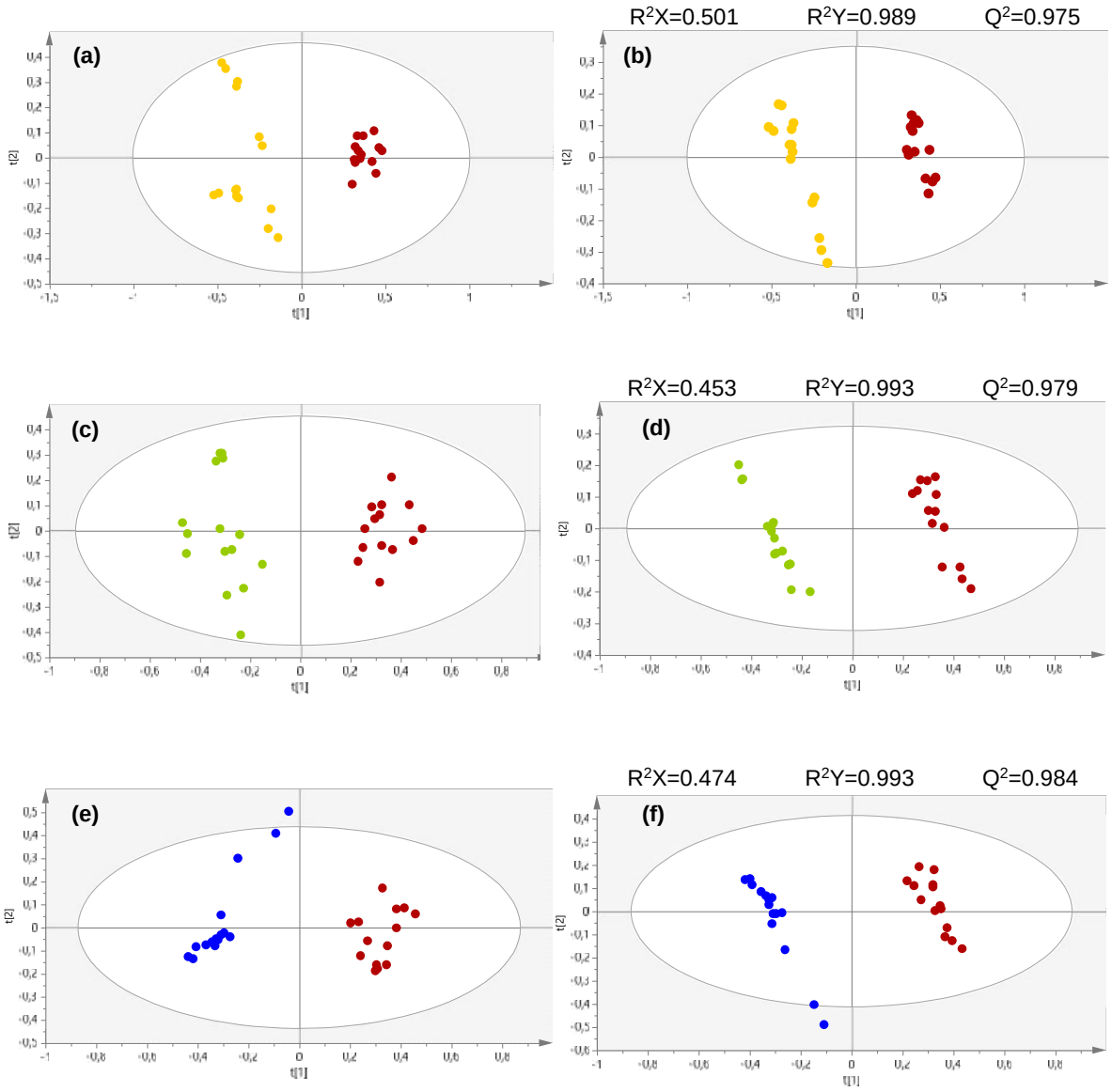

Figure 5. (a), (c) and (e) PCA-X and (b), (d) and (f) PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of the normal cell line (SV-HUC-1, n=14, ●) and the cancer cell lines 5637 (n=15, ●), J82 (n=15, ●) and Scaber (n=15, ●) extracellular medium (pH 7), respectively. The ellipses indicate the 95% confidence limit of the model.

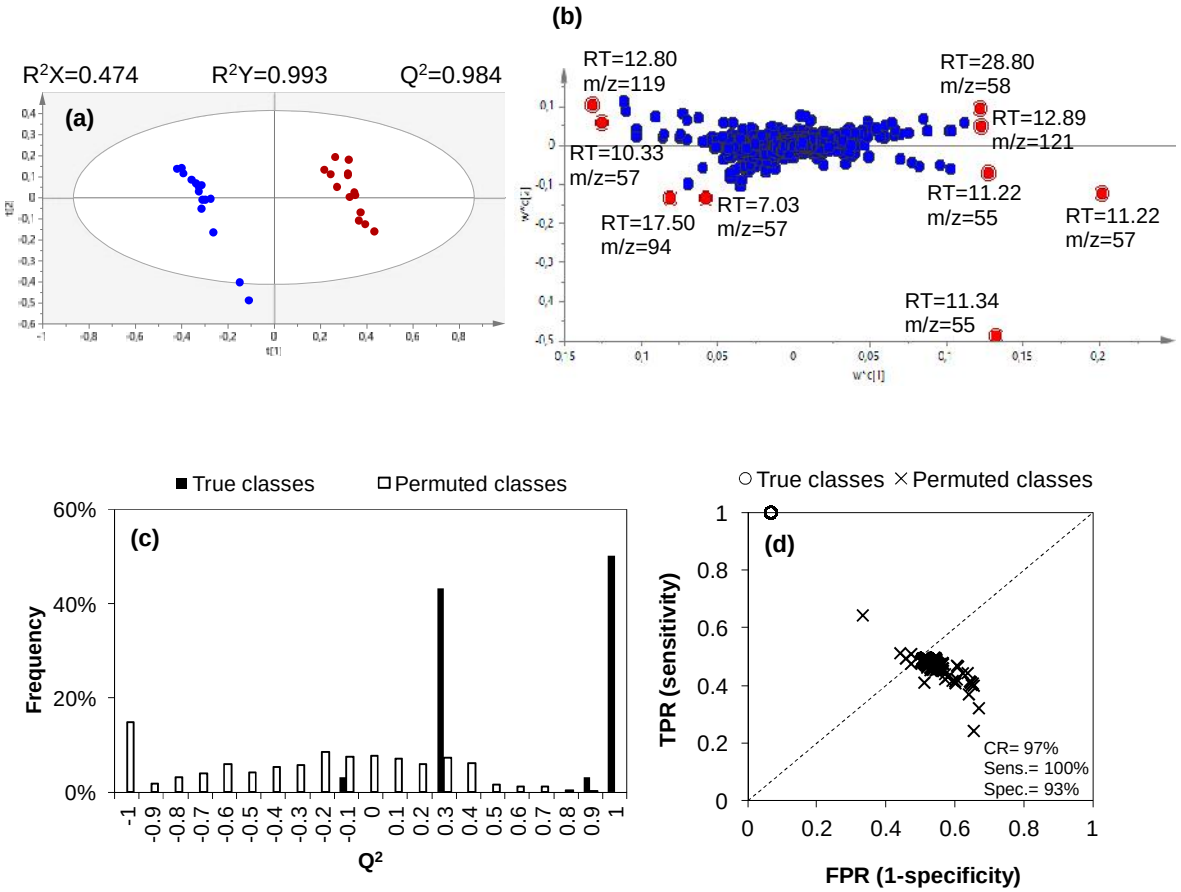

Figure 6. Example of the method used for the identification of the most significant metabolites represented by (a) PLS-DA score scatter plot obtained for the HS-SPME/GC-MS chromatograms of the normal cell line (SV-HUC-1, n=14, ●) and Scaber (n=15, ●) extracellular medium (pH 7) and its respective (b) loading scatter plot with RT and m/z of some of those metabolites. For the same group (Scaber vs SV-HUC-1), (c) Q^2 distribution and (d) ROC plot of true and permuted classes obtained by Monte Carlo cross validation are represented, as well.

PCA-X and PLS-DA models were also applied to compare the extracellular medium of the different cancer lines according to grade (figure 7) and to subtype (figure 8), and their respective R^2X, R^2Y and Q^2 values for PLS-DA models are also presented.

Regarding the overall models presented in figures 5b, 5d, 5f, 7b and 8b, MCCV results (Table 2) show that the prediction accuracy of the PLS-DA models was excellent with a classification rate ranging from 96% to 100% depending on the comparison. For instance, the PLS-DA model presented in figure 5c (Scaber vs SV-HUC-1) has a classification rate of 97%. Besides, calculation of permuted models enables the recognition of the groups separation observed in the PLS-DA models (figures 5, 7 and 8) as significant since 100% of the Q^2 values obtained by permutation were inferior than the original Q^2 distribution (example of figures 6c and 6d, Table 2).

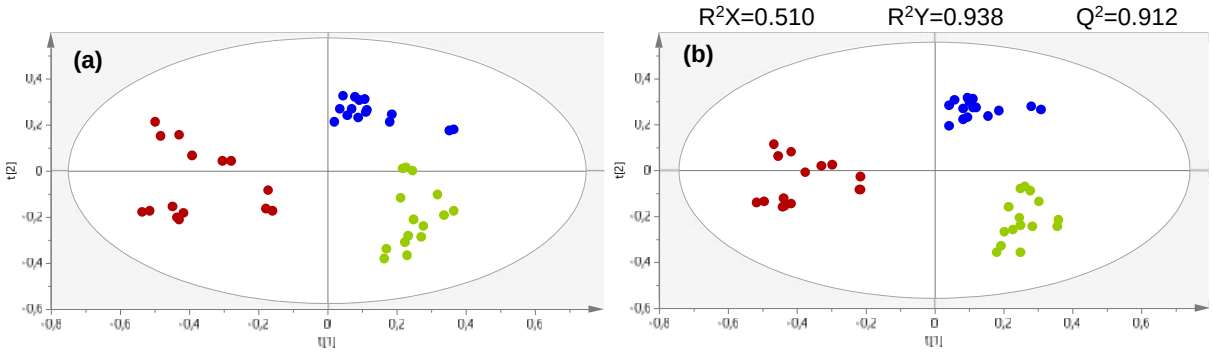

Figure 7. (a) PCA-X and (b) PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms that compare the LG cancer cell line 5637 (n=15, ●) with HG cancer cell lines J82 (n=15, ●) and Scaber (n=15, ●) extracellular medium (pH 7). The ellipses indicate the 95% confidence limit of the model.

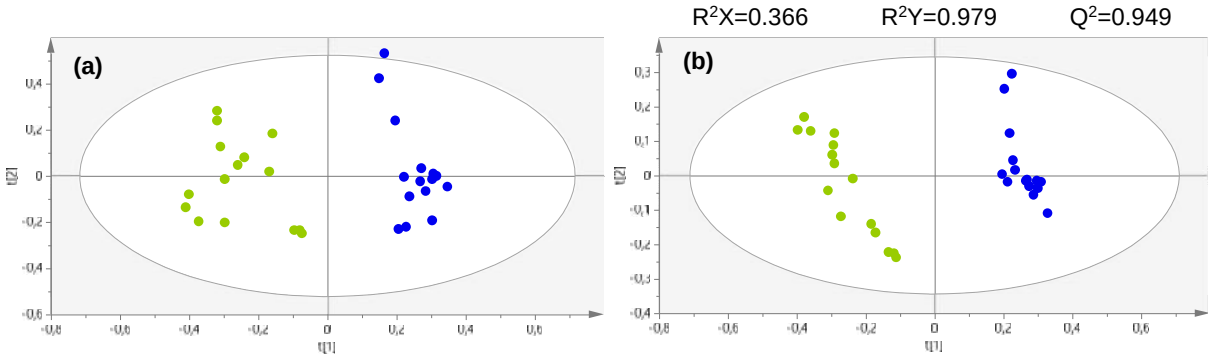

Figure 8. (a) PCA-X and (b) PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms that compare the HG TCC cell line J82 (n=15, ●) with the HG SCC cell line Scaber (n=15, ●) extracellular medium (pH 7). The ellipses indicate the 95% confidence limit of the model.

Table 2. MCCV parameters of true and permuted classes obtained for pH 7 when considering GC-MS full data.

Models	True classes					Permuted classes				
	LV	Q^2	CR (%)	Sens. (%)	Spec. (%)	LV	Q^2	CR (%)	Sens. (%)	Spec. (%)
GC-MS full data										
5637 vs SV-HUC-1	2	0.75	96	100	93	1	-0.22	47	47	47
J82 vs SV-HUC-1	2	0.69	97	100	93	1	-0.22	49	49	49
Scaber vs SV-HUC-1	3	0.91	97	100	93	1	-0.24	49	49	48
J82&Scaber vs 5637	5	0.95	100	100	100	1	-0.29	59	80	17
J82 vs Scaber	1	0.92	100	100	100	1	-0.18	49	49	49

Note: LV – number of latent variables, Q^2 – medium predictive power, CR – classification rate, sens. – sensitivity, spec. – specificity.

5.1.2. GC-MS analysis of the samples at pH 2

The GC-MS analysis of the samples at pH 2 consistently resulted in a total of 206 peaks, in which 107 were in common with pH 7 chromatographic peaks. Representative full scan chromatograms of each cell line are presented with some of the most abundant compounds, as well as the respective controls medium (figure 9).

Comparing the chromatograms from samples obtained at pH 7 and pH 2 (figures 3 and 9, respectively), it is possible to observe that altering the pH changed the VOC profile of samples. Acidification of samples generated more chromatographic peaks and metabolites such as 2-ethyl-1-hexanol and 4-methylbenzaldehyde became more intense at pH 2 compared to pH 7.

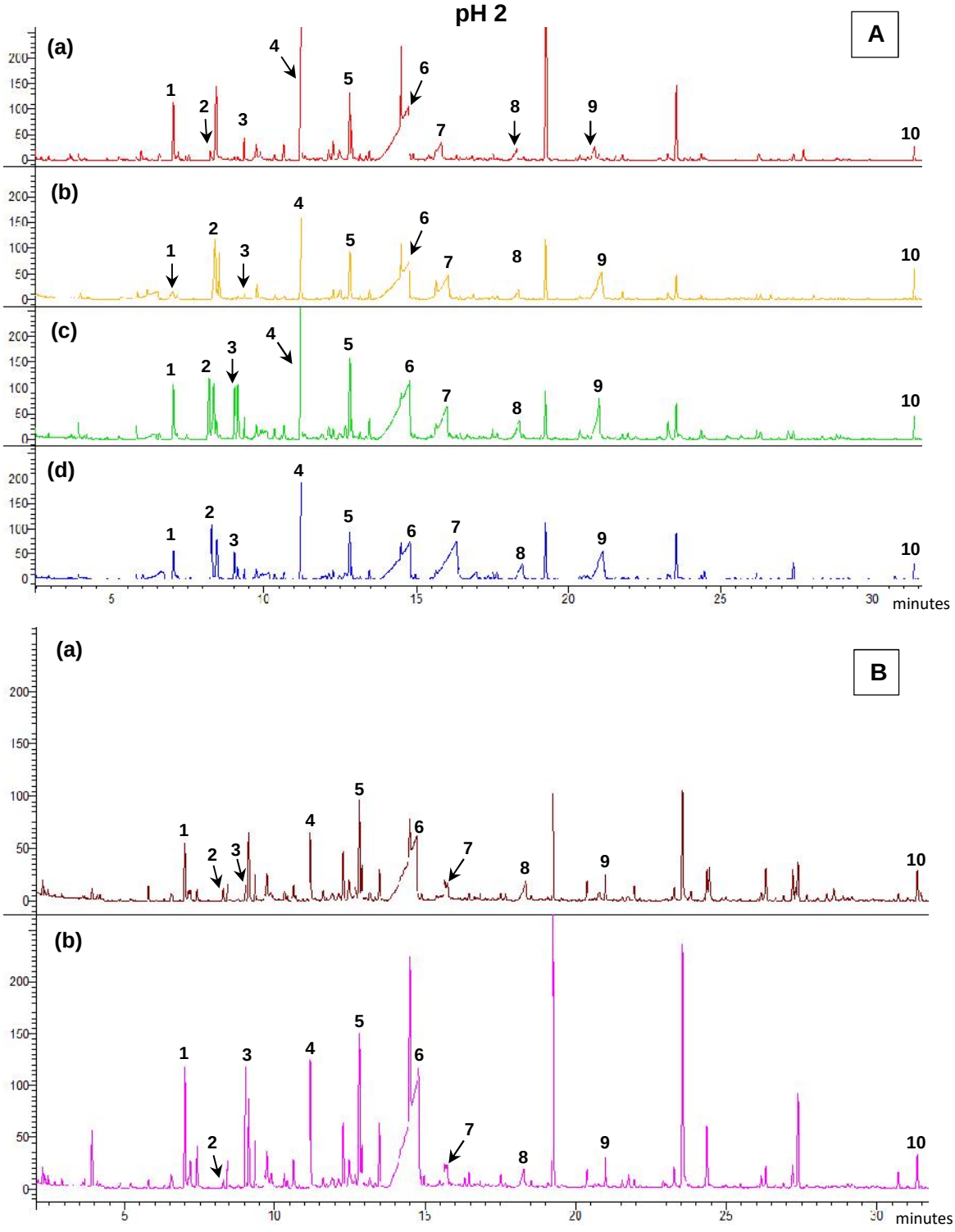

Figure 9. A. Representative full scan chromatograms obtained from the analysis of the extracellular medium of (a) SV-HUC-1, (b) 5637, (c) J82 and (d) Scaber, at pH 2. B. Representative full scan chromatograms obtained from the analysis of the controls medium of (a) SV-HUC-1 and (b) 5637, J82 and Scaber, at pH 2. Legend: 1. cyclohexanol; 2. 2,6-dimethylheptan-4-one; 3. 4-fluorobenzaldehyde (IS); 4. 2-ethyl-1-hexanol; 5. 4-methylbenzaldehyde; 6. 2-ethylhexanoic acid; 7. octanoic acid; 8. nonanoic acid; 9. decanoic acid; 10. hexadecanal. Note: the compounds are characterized by their IUPAC names.

After chemometric analysis, QCs at pH 2 were also closely clustered but slightly decentred in the PCA-X model (figure 10), compared to pH 7. As expected, a clear separation between the culture medium of the cell lines and the controls was observed on PC1, as well (figure 10).

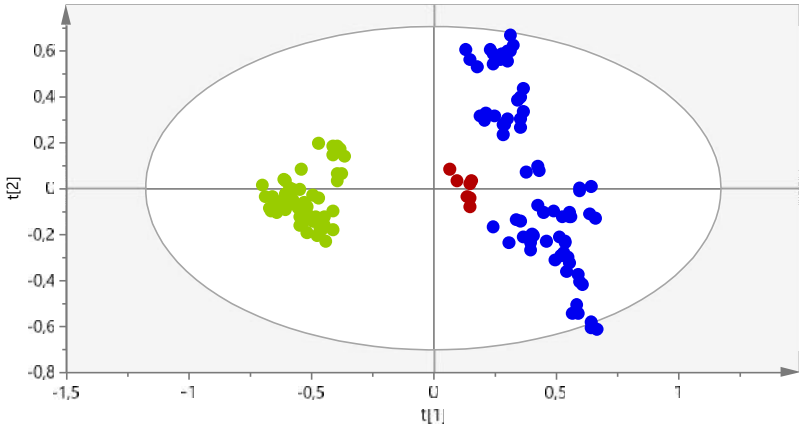

Figure 10. PCA-X score scatter plot obtained for the HS-SPME/GC-MS chromatograms of all samples at pH 2, namely medium of the controls (n=60, ●), QCs (n=8, ●) and extracellular medium of all cell lines (n=60, ●). The ellipses indicate the 95% confidence limit of the model.

Considering the extracellular medium of the cell lines, PCA-X and PLS-DA scores scatter plots were performed to see whether a clear separation was also evident as that observed between pH 7 samples. Indeed, PCA-X and PLS-DA models were similar as those obtained for pH 7 and a good separation between the extracellular medium of the BC cell lines and the normal cell line was observed (figures 11a to 11f). Values for R^2X, R^2Y and Q^2 are also presented in each figure. One sample from the cell line 5637 group was excluded from the PLS-DA model and from the following statistical analyses since it was considered an outlier for the same reasons as those considered for the SV-HUC-1 outlier (pH 7). The method used for the identification of the metabolites responsible for the groups separation in all comparisons, at pH 2, was identical as performed for pH 7 (VIP>1). In figures 12a and 12b, it is also illustrated an example of a PLS-DA model (Scaber vs SV-HUC-1) and its corresponding loading scatter plot, respectively. The metabolites' variations were also verified by using boxplots performed on Graphpad and by determining the % of variation and effect size values. The Q^2 distribution and ROC plot of true and permuted classes obtained by MCCV (Table 3) for the PLS-DA model represented in figure 12a are also presented in figures 12c and 12d, respectively.

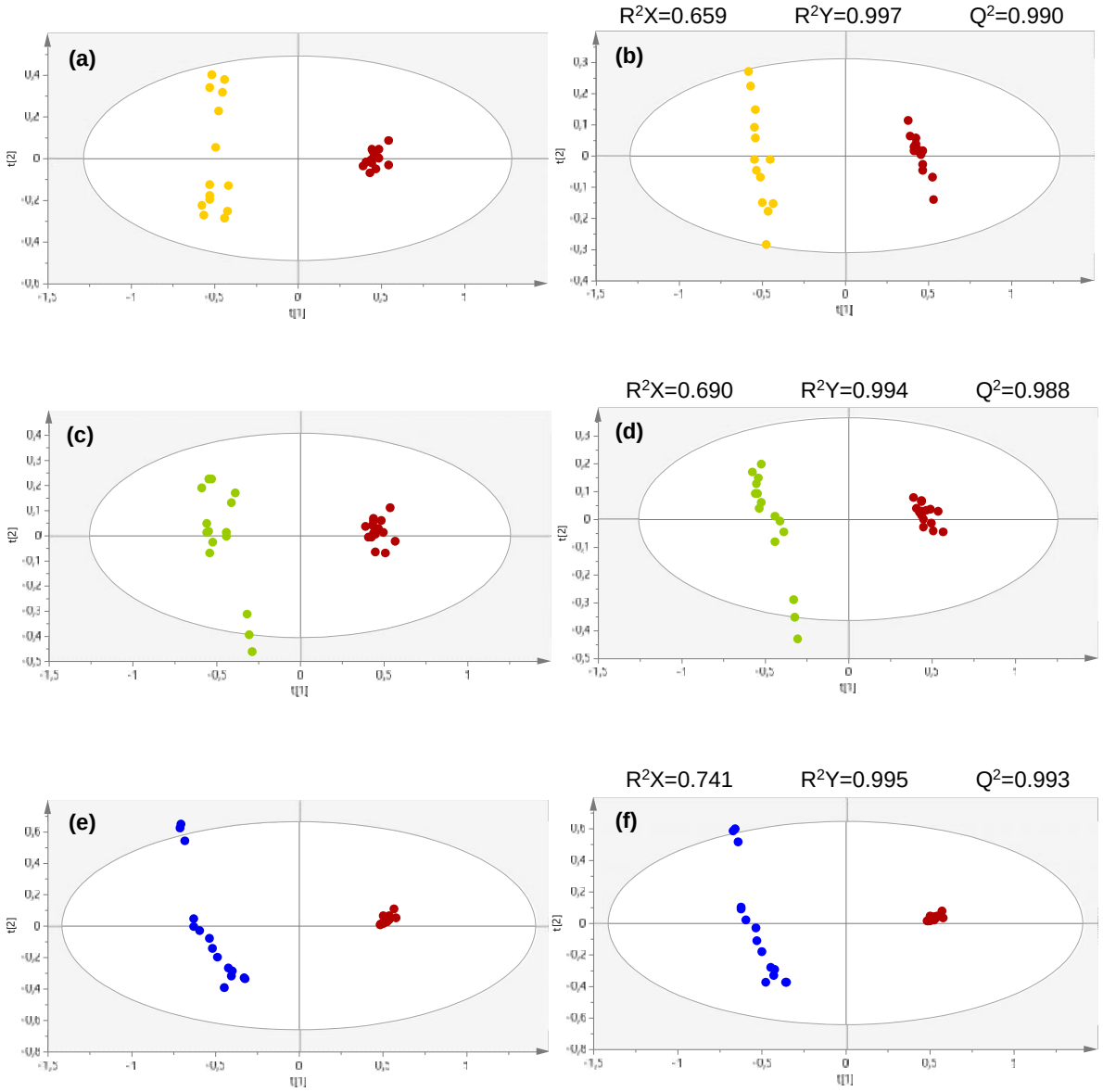

Figure 11. (a), (c) and (e) PCA-X and (b), (d) and (f) PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of the normal cell line (SV-HUC-1, n=15, ●) and the cancer cell lines 5637 (n=14, ●), J82 (n=15, ●) and Scaber (n=15, ●) extracellular medium (pH 2), respectively. The ellipses indicate the 95% confidence limit of the model.

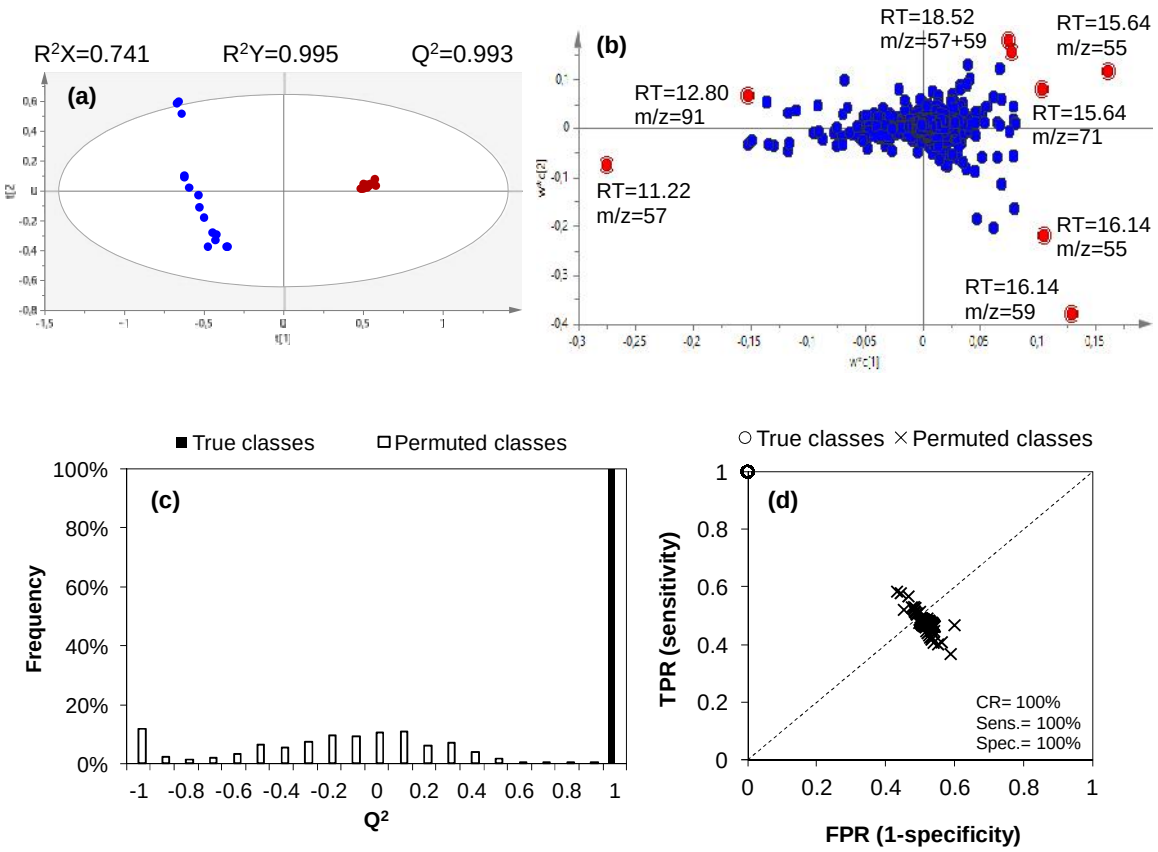

Figure 12. Example of the method used for the identification of the most significant metabolites represented by (a) PLS-DA score scatter plot obtained for the HS-SPME/GC-MS chromatograms of the normal cell line (SV-HUC-1, n=15, ●) and Scaber (n=15, ●) extracellular medium (pH 2) and its respective (b) loading scatter plot with RT and m/z of some of those metabolites. For the same group (Scaber vs SV-HUC-1), (c) Q^2 distribution and (d) ROC plot of true and permuted classes obtained by Monte Carlo cross validation are represented, as well.

PCA-X and PLS-DA models were also applied to compare the extracellular medium of the different cancer cell lines according to grade (figure 13) and to subtype (figure 14), at pH 2. The respective R^2X, R^2Y and Q^2 values for PLS-DA models are presented, as well.

Regarding the overall models presented in figures 11b, 11d, 11f, 13b and 14b, MCCV results (Table 3) show that the prediction accuracy of the PLS-DA models was excellent with an average classification rate of 100% for all comparisons, which is slightly better than those obtained for pH 7. In addition, calculation of permuted models at pH 2 also enables the recognition of the groups separations observed in the PLS-DA models (figures 11, 13 and 14) as significant since 100% of the Q^2 values obtained by permutation were, again, inferior to the original Q^2 distribution (example of figures 12c and 12d, Table 3).

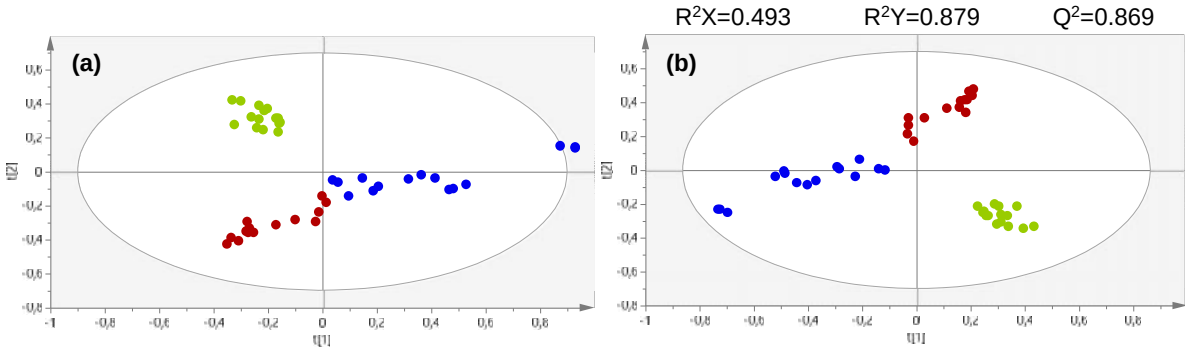

Figure 13. (a) PCA-X and (b) PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms that compare the LG cancer cell line 5637 (n=14, ●) with HG cancer cell lines J82 (n=15, ●) and Scaber (n=15, ●) extracellular medium (pH 2). The ellipses indicate the 95% confidence limit of the model.

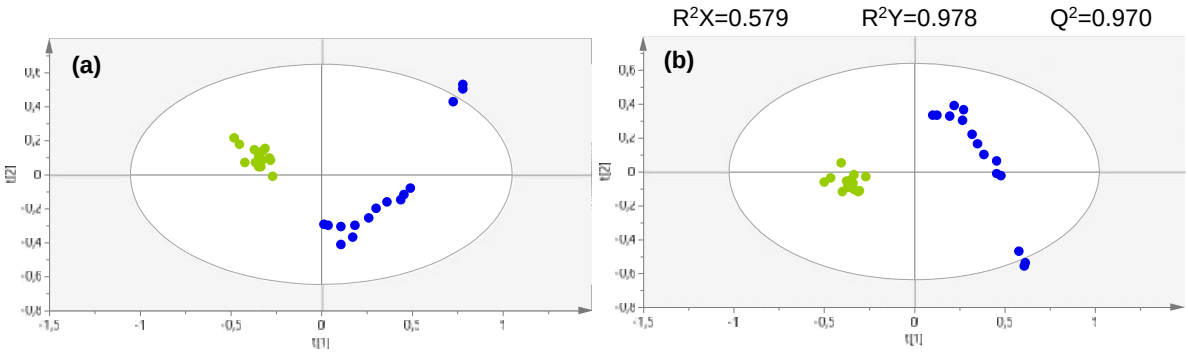

Figure 14. (a) PCA-X and (b) PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms that compare the HG TCC cell line J82 (n=15, ●) with HG SCC cell line Scaber (n=15, ●) extracellular medium (pH 2). The ellipses indicate the 95% confidence limit of the model.

Table 3. MCCV parameters of true and permuted classes obtained for pH 2 when considering GC-MS full data.

Models	True classes					Permuted classes				
	LV	Q^2	CR (%)	Sens. (%)	Spec. (%)	LV	Q^2	CR (%)	Sens. (%)	Spec. (%)
GC-MS full data										
5637 vs SV HUC-1	1	0.99	100	100	100	1	-0.19	46	42	50
J82 vs SV HUC-1	1	0.98	100	100	100	1	-0.20	49	49	49
Scaber vs SV HUC-1	4	0.98	100	100	100	1	-0.19	49	48	49
J82&Scaber vs 5637	4	0.95	100	100	100	1	-0.17	60	80	19
J82 vs Scaber	2	0.97	100	100	100	1	-0.23	50	49	50

Note: LV – number of latent variables, Q^2 – medium predictive power, CR – classification rate, sens. – sensitivity, spec. – specificity.

5.2. Untargeted approach: selection of the discriminative metabolites

After statistical analysis, several compounds from pH 7 samples analysis revealed to be significant in discriminating the cell lines culture medium, whereas, from the pH 2 analysis, the total of number of discriminant metabolites was lower, being most of them in common with pH 7 metabolites. The general characteristics of these metabolites, such as retention time (RT), characteristic ions (m/z), HMDB identification and matrices or cellular locations where they have been previously found are summarized in Supplementary Table S2 (Appendix) for both pHs.

5.2.1. Discriminative VOCs extracted at pH 7

5.2.1.1. Cancer versus normal cell lines extracellular medium

By studying the loading scatter plots and considering a VIP>1, it was possible to discover which compounds were responsible for the separation of each cancer cell line from the normal cell line extracellular medium. The analysis resulted in a total of 68 VOCs significantly altered (Table 4). Of those 68 VOCs, only 21 appeared altered in the three cancer cell lines, such as isopentanol, cyclohexanol, 4-methylheptan-2-one, benzaldehyde, 4-methylnonane, 2-nonanone, dodecane and tetradecane, which were increased in cancer cells medium, and 2-ethyl-1-hexanol, octanol and 2-phenylpropanol, which were decreased in cancer cells medium compared to SV-HUC-1 medium. However, 1,3-dimethylbenzene, acetophenone, 1-phenoxypropan-2-ol and 2-pentadecanone, despite being altered in the three cancer cells medium compared to the normal one, were observed to have a different trend of alteration among cancer cells. For instance, 1,3-dimethylbenzene and 1-phenoxypropan-2-ol were found increased and decreased, respectively, in 5637 and J82 compared to SV-HUC-1 culture medium, whereas in Scaber culture medium the variation is inverse, which may be related to the subtype of BC as 5637 and J82 are TCC and Scaber is SCC. Acetophenone was increased in 5637 but decreased in J82 and Scaber culture medium compared to the normal cells, which in this case, this difference between BC cells might be related to the grade of cancer. Interestingly, 2-pentadecanone was found decreased in 5637 and Scaber but augmented in J82 culture medium, which may be a characteristic of J82 cell line (high-grade TCC).

The remaining 47 metabolites were altered either in two of the cancer cells or altered in only one of the cancer cells. VOCs that only appeared significantly altered in 5637 and J82 culture medium compared with SV-HUC-1 culture medium may be an

indicative of the BC subtype TCC. Among those compounds, ethylbenzene, 6-methylheptan-2-one, 1,2,4-trimethylbenzene and two unidentified metabolites (unknowns 9 and 17) were increased and 2-tridecanone and unknown 7 were decreased in the extracellular medium of TCC cells. Additionally, 3-ethyloctane and nonanal were increased in 5637 but decreased in J82 compared to SV-HUC-1 culture medium, which can also correlate with low or high-grade TCC.

There were also some compounds that were found statistically different only in J82 and Scaber culture medium compared with the normal cell one, but were not significantly altered in 5637 culture medium, evidencing again the influence of cancer grade. Cyclohexanone, benzyl alcohol, menthol, α-terpineol, 2-hydroxy-2-methyl-1-phenylpropan-1-one and unknown 21 were diminished, whereas unknowns 4 and 16 were augmented in HG cancer cells culture medium (Table 4). One of the VOCs, phenol, was decreased in J82 but augmented in Scaber culture medium compared to SV-HUC-1, so the subtype of BC may be influencing the difference.

Concerning 5637 extracellular medium, methyl isobutyl ketone, styrene, 2,6-dimethylnonane and naphthalene were increased whereas the unknowns 13 and 22-24 were decreased compared to SV-HUC-1 extracellular medium. In turn, 1,2-dimethylbenzene, 2-undecanone and the unknown 19 were increased and 2-methyl-2-butanol was decreased in J82 extracellular medium compared to SV-HUC-1 extracellular medium (Table 4). As for Scaber extracellular medium, 2-pentanone, 2-methyl-2-pentanol, benzothiazole, dodecanal and 1-dodecanol were increased whereas the unknowns 10 and 26 were decreased compared to SV-HUC-1 extracellular medium (Table 4). These VOCs are important as they may be specific signatures of each cell line. Interestingly, 2-pentadecanone was the only compound that was not present in the controls medium, which suggests that its origin arises undoubtedly from the bladder cells metabolism.

Table 4. List of VOCs, from the analysis of samples prepared at pH 7, selected as important in discriminating the cancer cell lines (5637, J82 and Scaber) from the normal cell line SV-HUC-1, which are characterized by their IUPAC (and common) name. The % variation (± % uncertainty), ES, ES_{SE} and *p*-values are represented for each VOC.

Metabolite	5637 (n=15) vs SV HUC-1 (n=14)				J82 (n=15) vs SV HUC-1 (n=14)				Scaber (n=15) vs SV HUC-1 (n=14)			
		ES (±ES_{SE})[a]	% variation (± % uncertainty)	*p*-value[b]		ES (±ES_{SE})[a]	% variation (± % uncertainty)	*p*-value[b]		ES (±ES_{SE})[a]	% variation (± % uncertainty)	*p*-value[b]
2-methylbutan-2-ol (2-Methyl-2-butanol)					↓			$8.00x10^{-3}$				
2-Pentanone									↑	1.74 (0.82)	44.54 (7.42)	$1.10x10^{-3}$
2-Methyl-2-pentanol									↑	0.99 (0.74)	30.31 (9.48)	$1.10x10^{-2}$
3-methylbutan-1-ol (Isopentanol)	↓	-1.28 (0.77)	-21.05 (6.51)	$1.40x10^{-3}$	↓	-3.48 (1.12)	-51.02 (7.00)	$<1.00x10^{-4}$ *	↓	-3.91 (1.21)	-57.84 (7.39)	$<1.00x10^{-4}$ *
4-methylpentan-2-one (Methyl isobutyl ketone)	↑	2.65 (0.97)	67.63 (6.77)	$<1.00x10^{-4}$ *								
2,3-Dimethylhexane	↑	3.46 (1.12)	81.48 (5.94)	$<1.00x10^{-4}$ *					↑	2.26 (0.90)	55.90 (6.86)	$<1.00x10^{-4}$ *
4-methyl-3-penten-2-one	↑	2.36 (0.92)	40.01 (5.02)	$<1.00x10^{-4}$ *					↑	1.62 (0.81)	35.33 (6.61)	$3.00x10^{-4}$ *
Ethylbenzene	↑	1.39 (0.78)	281.66 (29.89)	$<1.00x10^{-4}$ *	↑	1.31 (0.77)	237.86 (29.36)	$1.00x10^{-4}$ *				
1,3-dimethylbenzene (m-Xylene)	↑	1.32 (0.77)	243.20 (29.60)	$<1.00x10^{-4}$ *	↑	1.25 (0.76)	205.09 (28.89)	$1.00x10^{-4}$ *	↓		-41.07 (31.98)	$3.00x10^{-2}$
Unknown 1	↑	1.53 (0.80)	61.38 (10.89)	$<1.00x10^{-4}$ *					↑	1.35 (0.78)	21.12 (5.02)	$<1.00x10^{-4}$ *
Cyclohexanol	↑	1.68 (0.82)	34.84 (6.28)	$<1.00x10^{-4}$ *	↑	0.89 (0.73)	18.06 (6.64)	$3.00x10^{-4}$ *	↑	0.74 (0.72)	14.88 (6.62)	$1.00x10^{-3}$ *
Ethenylbenzene (Styrene)	↑	1.33 (0.77)	50.62 (10.82)	$1.40x10^{-3}$								
1,2-dimethylbenzene (o-Xylene)					↑	1.37 (0.78)	338.25 (32.57)	$1.00x10^{-4}$ *				
Cyclohexanone					↓	-4.87 (1.41)	-87.43 (11.34)	$<1.00x10^{-4}$ *	↓		-11.70 (7.87)	$9.80x10^{-3}$
Unknown 3	↑	1.59 (0.80)	97.58 (14.62)	$<1.00x10^{-4}$ *	↑	1.23 (0.76)	59.46 (13.27)	$2.00x10^{-3}$	↑	4.07 (1.24)	179.10 (8.25)	$<1.00x10^{-4}$ *
4-methylheptan-2-one	↑	2.97 (1.03)	87.46 (7.27)	$<1.00x10^{-4}$ *	↑	2.55 (0.95)	57.19 (6.19)	$<1.00x10^{-4}$ *	↑	3.53 (1.13)	85.96 (6.06)	$<1.00x10^{-4}$ *
6-Methylheptan-2-one	↑	1.34 (0.77)	117.25 (19.55)	$2.00x10^{-4}$ *	↑	1.15 (0.75)	84.42 (18.35)	$1.00x10^{-4}$ *				
Benzaldehyde	↑	1.44 (0.79)	42.42 (8.64)	$<1.00x10^{-4}$ *	↑	1.55 (0.80)	75.47 (12.53)	$<1.00x10^{-4}$ *	↑	1.89 (0.85)	61.81 (8.86)	$<1.00x10^{-4}$ *
4-Methylnonane	↑	1.55 (0.80)	46.95 (8.70)	$<1.00x10^{-4}$ *	↑	1.58 (0.80)	44.75 (8.23)	$3.00x10^{-4}$ *	↑	2.22 (0.90)	53.47 (6.75)	$<1.00x10^{-4}$ *
3-Ethyloctane	↑	0.75 (0.72)	16.34 (7.16)	$3.00x10^{-2}$	↓			$2.00x10^{-3}$				
Phenol					↓	-5.51 (1.56)	-78.17 (8.28)	$<1.00x10^{-4}$ *	↑	1.54 (0.80)	40.67 (7.78)	$<1.00x10^{-4}$ *
Unknown 4					↑	2.90 (1.01)	54.87 (5.28)	$<1.00x10^{-4}$ *	↑	4.13 (1.26)	71.16 (4.51)	$<1.00x10^{-4}$ *

(cont.) Metabolite	5637 (n=15) vs SV HUC-1 (n=14)				J82 (n=15) vs SV HUC-1 (n=14)				Scaber (n=15) vs SV HUC-1 (n=14)			
		ES (±ES_{SE})[a]	% variation (± % uncertainty)	*p*-value[b]		ES (±ES_{SE})[a]	% variation (± % uncertainty)	*p*-value[b]		ES (±ES_{SE})[a]	% variation (± % uncertainty)	*p*-value[b]
1,2,4-Trimethylbenzene	↑		54.80 (29.01)	$3.00x10^{-2}$	↑		44.98 (25.96)	$4.40x10^{-2}$				
Unknown 5	↑	3.94 (1.22)	102.03 (6.09)	$<1.00x10^{-4}$ *	↑	2.68 (0.97)	60.54 (6.17)	$<1.00x10^{-4}$ *	↑	3.16 (1.06)	79.07 (6.37)	$<1.00x10^{-4}$ *
Unknown 6	↑	1.61 (0.81)	33.56 (6.34)	$<1.00x10^{-4}$ *					↑	0.91 (0.73)	18.99 (6.79)	$8.00x10^{-4}$ *
2,6-Dimethylnonane	↑	4.42 (1.32)	107.88 (5.64)	$<1.00x10^{-4}$ *								
2-ethylhexan-1-ol (2-Ethyl-1-hexanol)	↓	-1.38 (0.78)	-29.00 (8.71)	$<1.00x10^{-4}$ *	↓		-9.56 (7.45)	$1.00x10^{-4}$ *	↓			$7.00x10^{-4}$ *
Phenylmethanol (Benzyl alcohol)					↓		-48.24 (32.79)	$2.10x10^{-2}$	↓	-1.13 (0.75)	-78.94 (41.16)	$<1.00x10^{-4}$ *
Unknown 7	↓			$3.00x10^{-2}$	↓		-13.17 (7.26)	$5.00x10^{-4}$ *				
Unknown 8	↑	3.18 (1.06)	58.97 (5.09)	$<1.00x10^{-4}$ *	↑	2.01 (0.86)	34.12 (5.15)	$<1.00x10^{-4}$ *	↑	2.19 (0.89)	37.43 (5.11)	$<1.00x10^{-4}$ *
1-phenylethan-1-ol (1-Phenylethanol)	↓	-1.24 (0.76)	-24.75 (8.11)	$<1.00x10^{-4}$ *					↓	-1.53 (0.80)	-29.77 (8.13)	$<1.00x10^{-4}$ *
Unknown 9	↑	3.39 (1.10)	57.69 (4.70)	$<1.00x10^{-4}$ *	↑	3.54 (1.13)	37.37 (3.16)	$<1.00x10^{-4}$ *				
1-phenylethan-1-one (Acetophenone)	↑	0.94 (0.74)	21.16 (7.26)	$1.00x10^{-2}$	↓	-2.68 (0.97)	-55.68 (10.23)	$<1.00x10^{-4}$ *	↓	-1.24 (0.76)	-25.41 (8.36)	$<1.00x10^{-4}$ *
Octan-1-ol (Octanol)	↓	-3.41 (1.11)	-32.49 (4.04)	$<1.00x10^{-4}$ *	↓	-1.22 (0.76)	-16.41 (5.21)	$4.90x10^{-3}$	↓	-3.52 (1.13)	-34.04 (4.13)	$<1.00x10^{-4}$ *
Unknown 10									↓	-0.88 (0.73)	-11.91 (5.13)	$1.20x10^{-3}$
Unknown 11	↑	5.20 (1.49)	144.58 (5.74)	$<1.00x10^{-4}$ *	↑	2.34 (0.91)	74.97 (8.29)	$<1.00x10^{-4}$ *	↑	2.41 (0.93)	78.04 (8.26)	$<1.00x10^{-4}$ *
4-Methylbenzaldehyde	↑	1.34 (0.77)	45.26 (9.78)	$3.00x10^{-4}$ *					↑	1.94 (0.85)	44.42 (6.66)	$<1.00x10^{-4}$ *
2-phenylpropan-2-ol	↓	-0.76 (0.72)	-15.27 (7.74)	$<1.00x10^{-4}$ *	↓	-1.63 (0.81)	-32.08 (8.35)	$<1.00x10^{-4}$ *	↓	-1.37 (0.78)	-27.12 (8.13)	$<1.00x10^{-4}$ *
Unknown 12	↑	1.35 (0.78)	37.67 (8.33)	$4.00x10^{-4}$ *					↑	1.10 (0.75)	31.23 (8.70)	$5.00x10^{-3}$
2-Nonanone	↑	1.96 (0.86)	80.34 (10.38)	$<1.00x10^{-4}$ *	↑	2.17 (0.89)	81.01 (9.43)	$<1.00x10^{-4}$ *	↑	2.08 (0.87)	102.78 (11.58)	$<1.00x10^{-4}$ *
Methyl benzoate	↑	1.09 (0.75)	37.96 (10.44)	$1.00x10^{-2}$					↑	1.41 (0.78)	40.48 (8.50)	$3.00x10^{-4}$ *
Nonanal	↑	1.94 (0.85)	104.66 (12.56)	$<1.00x10^{-4}$ *	↓	-0.78 (0.72)	-36.80 (20.57)	$1.50x10^{-3}$ *				
Unknown 13	↓	-1.78 (0.83)	-40.21 (10.04)	$5.00x10^{-4}$ *								
Unknown 14	↑	3.12 (1.05)	42.27 (3.98)	$<1.00x10^{-4}$ *	↑	3.71 (1.17)	43.89 (3.45)	$<1.00x10^{-4}$ *	↑	2.03 (0.87)	36.43 (5.39)	$<1.00x10^{-4}$ *
(1R,2S,5R)-5-methyl-2-(propan-2-yl) cyclo-hexan-1-ol (Menthol)					↓	-0.96 (0.74)	-34.87 (15.62)	$1.90x10^{-2}$	↓	-1.07 (0.75)	-39.52 (16.37)	$1.00x10^{-3}$ *
Naphthalene	↑	2.22 (0.89)	41.21 (5.47)	$<1.00x10^{-4}$ *								
2-(1R)-4-methylcyclo-hex-3-en-1-ylpropan-2-ol (α-Terpineol)					↓	-0.98 (0.74)	-13.25 (5.13)	$<1.00x10^{-4}$ *	↓	-1.86 (0.84)	-24.65 (5.38)	$<1.00x10^{-4}$ *

(cont.) Metabolite	5637 (n=15) vs SV HUC-1 (n=14)			J82 (n=15) vs SV HUC-1 (n=14)			Scaber (n=15) vs SV HUC-1 (n=14)		
	ES ($\pm ES_{SE}$)[a]	% variation (± % uncertainty)	*p*-value[b]	ES ($\pm ES_{SE}$)[a]	% variation (± % uncertainty)	*p*-value[b]	ES ($\pm ES_{SE}$)[a]	% variation (± % uncertainty)	*p*-value[b]
Dodecane (N-Dodecane)	↑ 5.57 (1.57)	100.67 (4.27)	$<1.00x10^{-4}$ *	↑ 4.97 (1.44)	77.61 (4.00)	$<1.00x10^{-4}$ *	↑ 3.78 (1.18)	65.38 (4.63)	$<1.00x10^{-4}$ *
1,3-benzothiazole (Benzothiazole)							↑ 2.42 (0.93)	37.24 (4.62)	$<1.00x10^{-4}$ *
1-phenoxypropan-2-ol	↓	-11.82 (8.10)	$1.00x10^{-3}$	↓	-7.37 (8.09)	$2.60x10^{-2}$	↑ 1.57 (0.80)	45.83 (8.44)	$<1.00x10^{-4}$ *
Unknown 15	↑ 3.27 (1.08)	100.45 (7.27)	$<1.00x10^{-4}$ *	↑ 2.32 (0.91)	64.91 (7.51)	$<1.00x10^{-4}$ *	↑ 2.08 (0.87)	70.02 (8.84)	$<1.00x10^{-4}$ *
2-hydroxy-2-methyl-1-phenylpropan-1-one				↓ -0.85 (0.73)	-13.02 (5.83)	$<1.00x10^{-4}$ *	↓ -1.09 (0.75)	-16.71 (5.96)	$<1.00x10^{-4}$ *
Undecan-2-one (2-Undecanone)				↑ 1.03 (0.74)	62.76 (16.48)	$<1.00x10^{-4}$ *			
Unknown 16				↑	16.22 (9.49)	$<1.00x10^{-4}$ *	↑ 1.18 (0.76)	36.25 (9.23)	$<1.00x10^{-4}$ *
Unknown 17	↑	33.48 (19.25)	$<1.00x10^{-4}$ *	↑	27.18 (20.66)	$<1.00x10^{-4}$ *			
Unknown 19				↑ 1.33 (0.77)	82.12 (15.53)	$6.00x10^{-4}$ *			
Tetradecane	↑ 1.85 (0.84)	56.13 (8.40)	$<1.00x10^{-4}$ *	↑ 1.33 (0.77)	38.10 (8.53)	$<1.00x10^{-4}$ *	↑ 1.18 (0.76)	33.77 (8.71)	$<1.00x10^{-4}$ *
Dodecanal (Lauric aldehyde)							↑ 1.44 (0.79)	59.34 (11.31)	$1.00x10^{-4}$ *
Dodecan-1-ol (1-Dodecanol)							↑ 2.41 (0.93)	103.35 (10.04)	$<1.00x10^{-4}$ *
Unknown 20	↓ -1.37 (0.78)	-86.06 (39.18)	$<1.00x10^{-4}$ *				↓ -1.34 (0.77)	-83.46 (38.07)	$<1.00x10^{-4}$ *
Unknown 21				↓ -5.05 (1.45)	-92.55 (12.13)	$<1.00x10^{-4}$ *	↓ -4.21 (1.27)	-54.00 (6.24)	$<1.00x10^{-4}$ *
Tridecan-2-one (2-Tridecanone)	↓ -4.75 (1.39)	-84.50 (10.94)	$<1.00x10^{-4}$ *	↓ -5.03 (1.45)	-89.15 (11.35)	$<1.00x10^{-4}$ *			
Unknown 22	↓ -1.02 (0.74)	-16.22 (6.15)	$<1.00x10^{-4}$ *						
Pentadecan-2-one (2-Pentadecanone)	↓ -1.00 (0.74)	-20.52 (8.14)	$<1.00x10^{-4}$ *	↑ 4.15 (1.26)	86.46 (5.16)	$<1.00x10^{-4}$ *	↓ -3.12 (1.05)	-58.79 (9.47)	$<1.00x10^{-4}$ *
Unknown 23	↓ -2.32 (0.91)	-37.78 (7.14)	$<1.00x10^{-4}$ *						
Unknown 24	↓ -1.78 (0.83)	-29.65 (6.93)	$<1.00x10^{-4}$ *						
Unknown 25	↓ -1.12 (0.75)	-20.50 (7.23)	$<1.00x10^{-4}$ *				↑ 1.13 (0.75)	19.98 (5.71)	$<1.00x10^{-4}$ *
Unknown 26							↓ -0.90 (0.73)	-16.49 (7.07)	$<1.00x10^{-4}$ *

Note: ES, effect size;

↑, metabolites that are increased; ↓, metabolites that are decreased in the extracellular medium

[a]ES determined as described in reference 104; values in brackets correspond to high uncertainties; [b]95% significance level (*p*-value <0.05).

*Metabolites remaining significant after Bonferroni correction, with cut off *p*-value of $1.04x10^{-3}$ for 5637 vs SV-HUC-1 (0.05 divided by 48 metabolites), $1.16x10^{-3}$ for J82 vs SV-HUC-1 (0.05 divided by 43 metabolites) and $1.06x10^{-3}$ for Scaber vs SV-HUC-1 (0.05 divided by 47 metabolites).

The PCA-X and PLS-DA scores scatter plots that were performed considering the metabolites identified as significantly discriminative remained almost unchanged, with a clear separation between cancer cell lines and the normal cell line (figures 15a to 15f). R^2X, R^2Y and Q^2 values became slightly better as well (except for Scaber *versus* SV-HUC-1). Besides, MCCV results (Table 5) of the PLS-DA models in figures 15b, d and f confirm the robustness of the metabolic signature for each cancer cell line compared with the normal one, since the classification rate is 100% for all comparisons and Q^2 values obtained for the permuted classes are lower than the original Q^2 values.

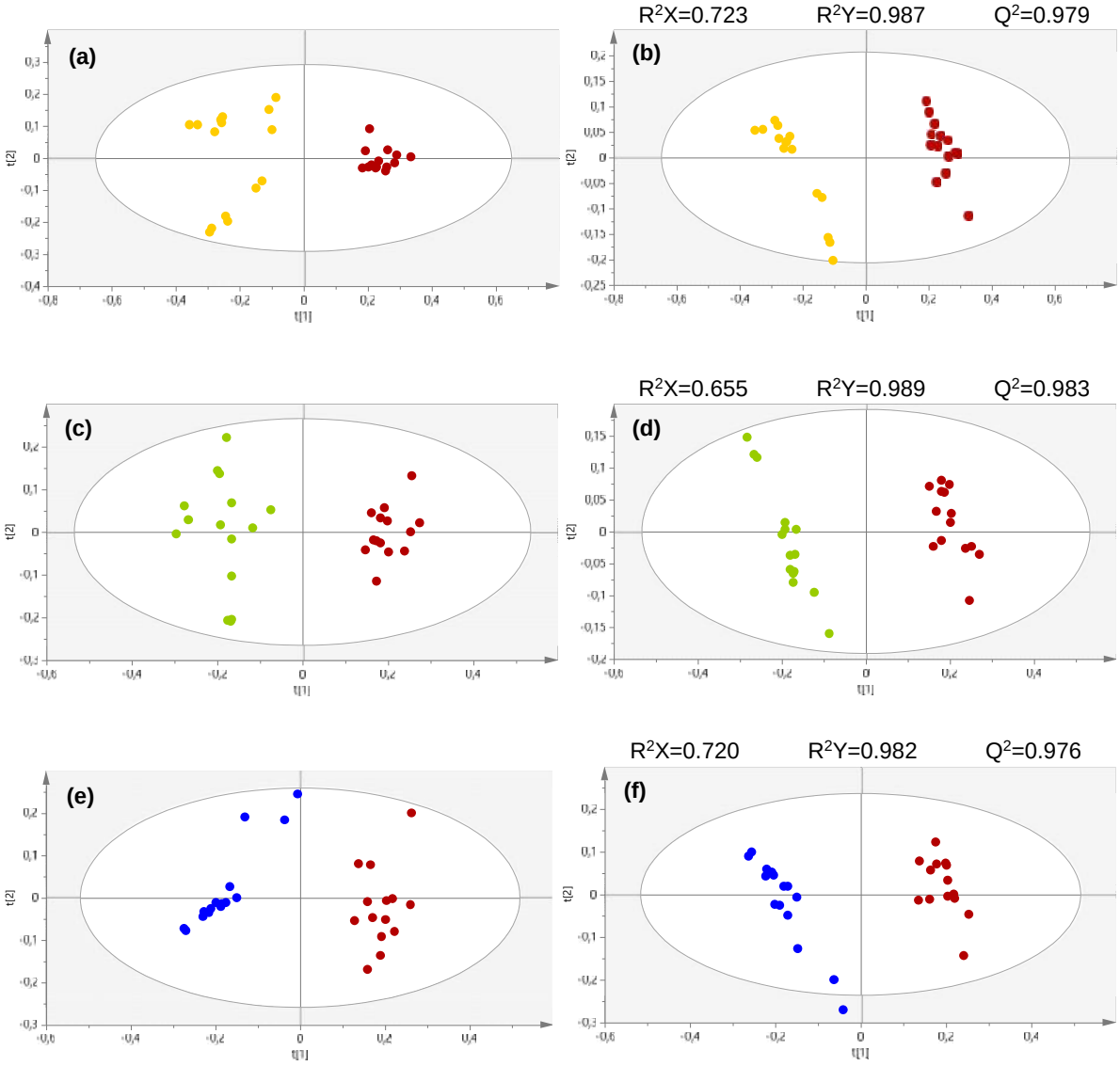

Figure 15. (a), (c) and (e) PCA-X and (b), (d) and (f) PLS-DA scores scatter plots obtained for the normal cell line (SV-HUC-1, n=14, ●) and the cancer cell lines 5637 (n=15, ●), J82 (n=15, ●) and Scaber (n=15, ●) extracellular medium, respectively, considering the discriminative metabolites only, at pH 7. The ellipses indicate the 95% confidence limit of the model.

Boxplots of some of the most important metabolites responsible for the separation of the cancer cell lines and SV-HUC-1 are also represented (figure 16).

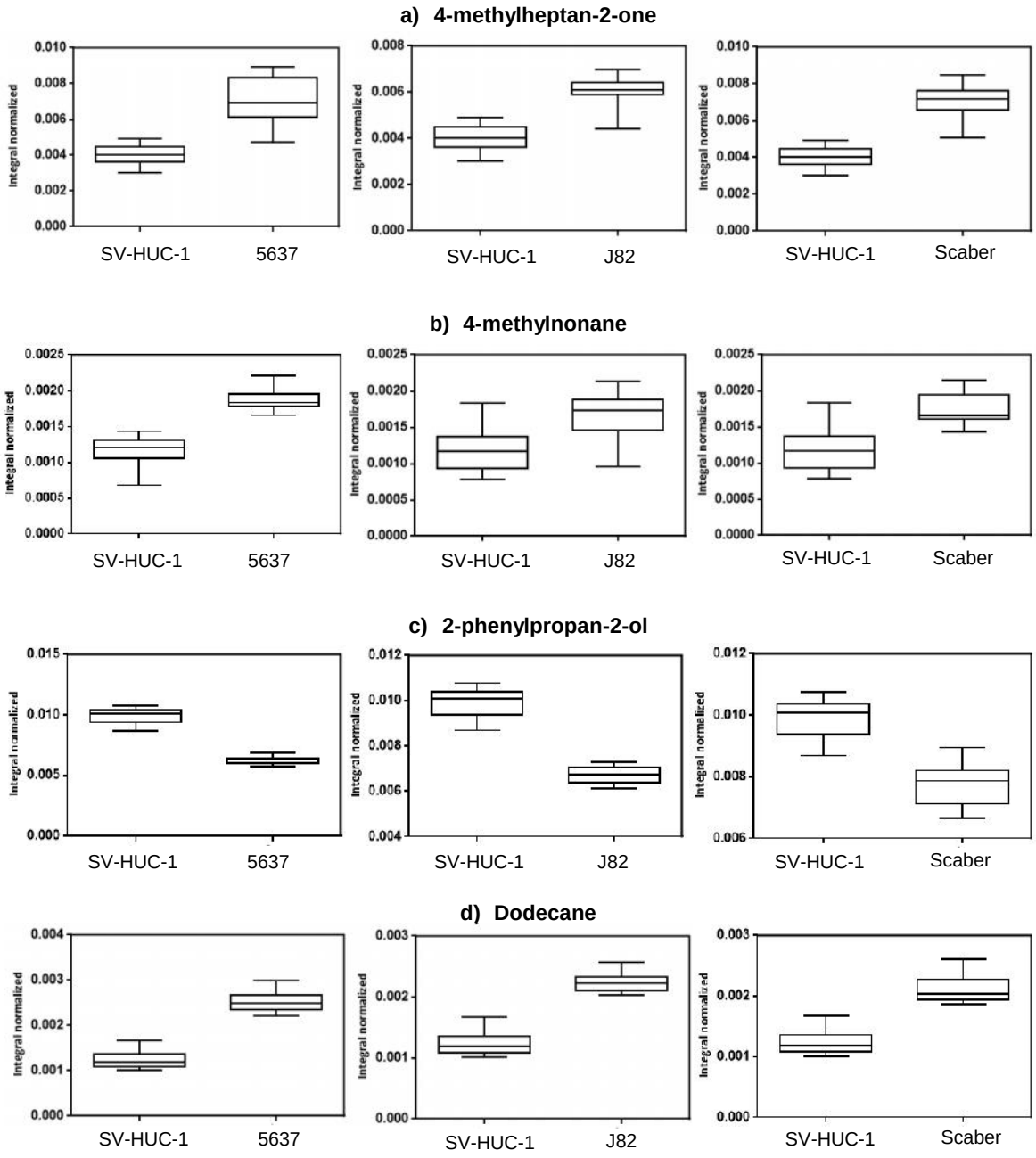

Figure 16. Boxplots for four VOCs found significantly altered between the normal cell line SV-HUC-1 (n=14) and the cancer cell lines 5637 (n=15), J82 (n=15) and Scaber (n=15) extracellular medium. *p*-value for a) 4-methylheptan-2-one; b) 4-methylnonane; (c) 2-phenylpropan-2-ol and (d) dodecane is $<1.00 \times 10^{-4}$. Note: *p*-values are equal in all combinations.

5.2.1.2. 5637 versus J82 and Scaber extracellular medium

A comparison between the low-grade (LG) cancer line 5637 and the high-grade (HG) BC cell lines J82 and Scaber was also performed to investigate which metabolites would enable the evaluation of tumour aggressiveness. A great number of metabolites appeared as being significantly altered, such as isopentanol, cyclohexanone, benzyl alcohol, α-terpineol, tetradecane, which were decreased in HG cancer cells medium, and benzaldehyde, 2-ethyl-1-hexanol, octanol, 1-phenoxypropan-2-ol and γ-nonalactone, which were augmented in HG cancer cells compared to 5637 culture medium, among others (Table 6).

The metabolic signature found for LG BC compared to HG BC cell lines was also confirmed by performing a PCA-X and PLS-DA scores scatter plots with the most important metabolites (figure 17). A clear separation between the cancer lines remained, as well as an excellent Q^2 value. In table 5, it is possible to see that MCCV confirm the robustness of the metabolic signature responsible for separating 5637 from HG cancer cells (classification rate of 100% and Q^2 values of permuted classes inferior to the original Q^2 values).

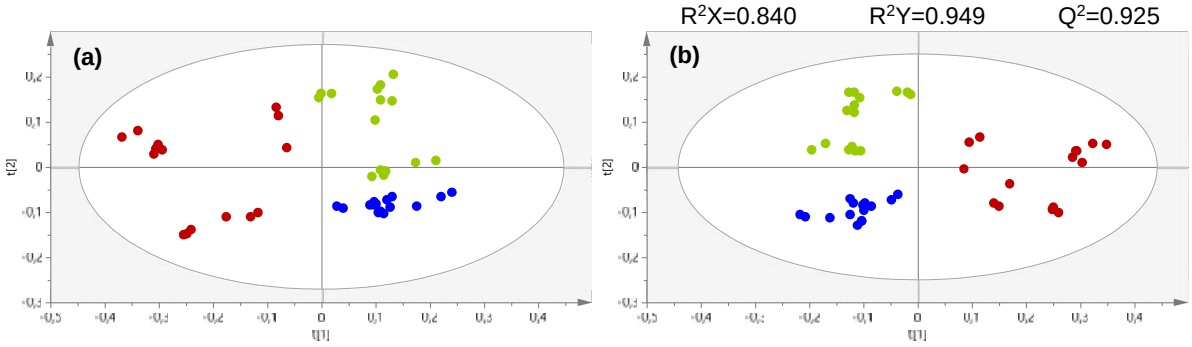

Figure 17. (a) PCA-X and (b) PLS-DA scores scatter plots obtained for the cancer cell lines 5637 (n=15, ●), J82 (n=15, ●) and Scaber (n=15, ●) extracellular medium, considering the discriminative metabolites only, at pH 7. The ellipses indicate the 95% confidence limit of the model.

Boxplots of some of the relevant metabolites in the separation of the cancer cell lines according to their grade are also represented (figure 18).

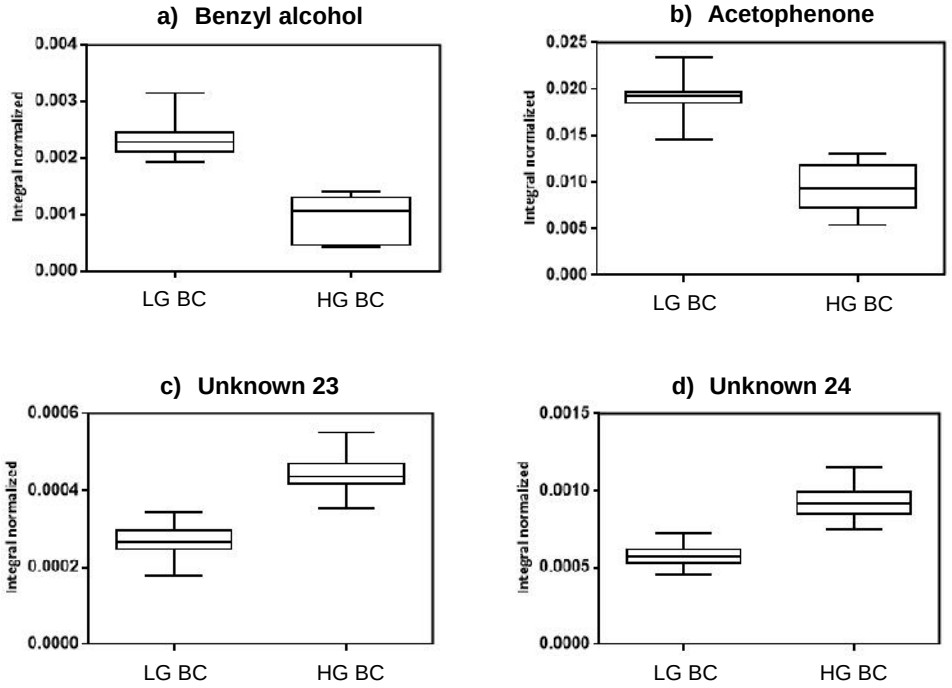

Figure 18. Boxplots for four VOCs found significantly altered between the LG cancer cell line 5637 (n=15) and the HG cancer cell lines J82 (n=15) and Scaber (n=15) extracellular medium. *p*-value for a) benzyl alcohol; b) acetophenone; c) unknown 23 and (d) unknown 24 is $<1.00 \times 10^{-4}$.

5.2.1.3. J82 versus Scaber extracellular medium

The two HG BC cell lines J82 and Scaber represent different subtypes of BC (TCC and SCC, respectively). Thus, we thought it would be worthy to compare these two cell lines and to study the metabolites that are significantly different between them. Indeed, it was found that 22 metabolites were also significantly altered between J82 and Scaber (Table 6), demonstrating that different histological subtypes of BC present different VOC profiles.

After identifying the discriminative metabolites, PCA-X and PLS-DA scores scatter plots were performed to confirm that those metabolites were indeed responsible for the separation of J82 and Scaber. As expected, the separation between the two cancer cell lines remained excellent (figure 19) and that is also demonstrated by the MCCV results in Table 5 (classification rate of 100% and Q^2 values of permuted classes lower than the original Q^2 values).

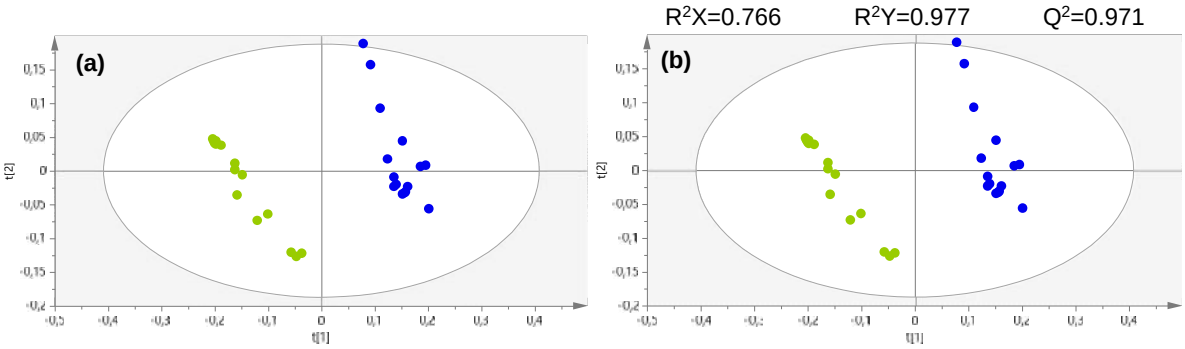

Figure 19. (a) PCA-X and (b) PLS-DA scores scatter plots obtained for the cancer cell lines J82 (n=15, ●) and Scaber (n=15, ●) extracellular medium, considering the discriminative metabolites only, at pH 7. The ellipses indicate the 95% confidence limit of the model.

Boxplots of three of the relevant metabolites in the separation of the two different BC subtypes cell lines are also represented (figure 20).

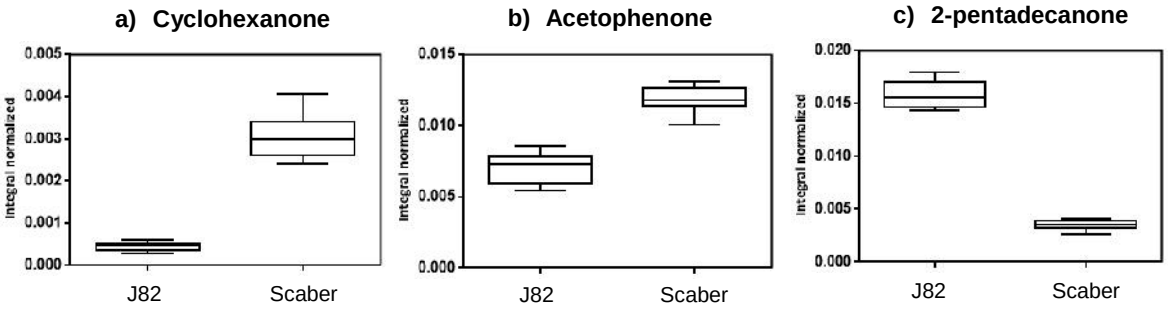

Figure 20. Boxplots for three VOCs found significantly altered between the cancer cell lines J82 (HG TCC, n=15) and Scaber (HG SCC, n=15) extracellular medium. *p*-value for a) cyclohexanone; b) acetophenone and c) 2-pentadecanone is $<1.00 \times 10^{-4}$.

Table 5. MCCV parameters of true and permuted classes obtained for pH 7 when considering the set of discriminant VOCs with statistical relevance (*p*-value < 0.05).

Models	True classes					Permuted classes				
	LV	Q^2	CR (%)	Sens. (%)	Spec. (%)	LV	Q^2	CR (%)	Sens. (%)	Spec. (%)
Set of discriminant VOCs										
5637 vs SV-HUC-1	3	0.99	100	100	100	1	-0.19	47	48	49
J82 vs SV-HUC-1	3	0.99	100	100	100	1	-0.22	48	51	45
Scaber vs SV-HUC-1	1	0.97	97	100	93	1	-0.24	46	46	46
J82&Scaber vs 5637	4	0.97	100	100	100	1	-0.28	57	71	29
J82 vs Scaber	1	0.92	100	100	100	1	-0.16	49	49	50

Note: LV – number of latent variables, Q^2 – medium predictive power, CR – classification rate, sens. – sensitivity, spec. – specificity.

Table 6. List of VOCs, from the analysis of samples prepared at pH 7, selected as important in discriminating between the cancer cell lines (5637 vs J82 vs Scaber), characterized by their IUPAC (and common) name. The % variation (± % uncertainty), ES, ES_{SE} and *p*-values are also represented for each VOC.

Metabolite	HG BC (J82 and Scaber) (n=30) vs LG BC (5637) (n=15)				J82 (n=15) vs Scaber (n=15)			
	ES ($\pm ES_{SE}$)[a]		% variation (± % uncertainty)	*p*-value[b]	ES ($\pm ES_{SE}$)[a]		% variation (± % uncertainty)	*p*-value[b]
2-methylbutan-2-ol (2-Methyl-2-butanol)	↓		-8.14 (3.97)	$4.00x10^{-2}$	↓	-1.26 (0.77)	-15.60 (4.77)	$3.00x10^{-3}$
2-Methyl-2-pentanol	↓		-13.87 (8.31)	$2.00x10^{-2}$	↓	-1.17 (0.76)	-22.14 (7.56)	$6.00x10^{-3}$
3-methylbutan-1-ol (Isopentanol)	↓	-1.60 (0.69)	-15.43 (3.13)	$<1.00x10^{-4}$ *				
4-methylpentan-2-one (Methyl isobutyl ketone)	↓	-1.52 (0.68)	-24.82 (6.50)	$2.00x10^{-4}$ *				
Butyl acetate (n-Butyl acetate)					↓	-1.53 (0.80)	-50.00 (15.50)	$<1.00x10^{-4}$ *
2-methylpentan-1-ol	↓	-1.49 (0.68)	-31.93 (8.42)	$<1.00x10^{-4}$ *				
Ethylbenzene	↓	-0.77 (0.73)	-46.28 (28.99)	$1.30x10^{-2}$	↑	1.62 (0.81)	368.23 (28.36)	$3.00x10^{-3}$
1,3-dimethylbenzene (m-Xylene)	↓	-0.77 (0.73)	-46.94 (29.56)	$6.00x10^{-4}$ *	↑	1.66 (0.81)	416.36 (28.96)	$5.00x10^{-4}$ *
Unknown 1	↓	-1.44 (0.68)	-27.90 (10.27)	$<1.00x10^{-4}$ *	↓	-1.04 (0.74)	-7.87 (2.79)	$8.00x10^{-3}$
Cyclohexanol	↓	-2.17 (0.76)	-13.63 (2.34)	$<1.00x10^{-4}$ *				
Ethenylbenzene (Styrene)	↓	-1.39 (0.67)	-29.44 (9.84)	$3.00x10^{-4}$ *	↑	1.38 (0.78)	29.67 (6.64)	$1.00x10^{-3}$ *
1,2-dimethylbenzene (o-Xylene)	↓	-0.72 (0.63)	-47.88 (32.73)	>0.05	↑	1.63 (0.81)	585.21 (32.45)	$1.00x10^{-2}$
Cyclohexanone	↓	-1.89 (0.72)	-55,88 (11.14)	$<1.00x10^{-4}$ *	↓	-6.88 (1.88)	-85.76 (7.75)	$<1.00x10^{-4}$ *
Unknown 2					↓	-1.59 (0.80)	-32.98 (8.82)	$4.00x10^{-4}$ *
Unknown 3					↓	-2.25 (0.90)	-42.87 (8.62)	$<1.00x10^{-4}$ *
4-methylheptan-2-one					↓	-1.39 (0.78)	-15.47 (4.30)	$4.00x10^{-4}$ *
Benzaldehyde	↑	0.73 (0.63)	18.41 (5.21)	$1.00x10^{-3}$ *				
4-Methylnonane	↓	-2.85 (0.85)	-29.65 (3.65)	$<1.00x10^{-4}$ *				
3-Ethyloctane	↓	-1.78 (0.71)	-15.36 (3.10)	$<1.00x10^{-4}$ *	↓	-0.82 (0.73)	-7.24 (3.23)	$2.00x10^{-2}$
Phenol					↓	-5.02 (1.45)	-84.48 (10.35)	$<1.00x10^{-4}$ *
Unknown 4	↓	-1.07 (0.65)	-11.45 (3.38)	$1.00x10^{-3}$ *				
Unknown 5	↓	-1.60 (0.69)	-15.45 (3.61)	$<1.00x10^{-4}$ *	↓	-1.10 80.75)	-10.35 (3.53)	$8.00x10^{-4}$ *
Unknown 6	↓	-1.96 (0.73)	-14.04 (2.40)	$<1.00x10^{-4}$ *	↓	-0.95 80.74)	-7.04 (2.73)	$2.00x10^{-2}$
2-ethylhexan-1-ol (2-Ethyl-1-hexanol)	↑	2.39 (0.78)	30.00 (3.78)	$<1.00x10^{-4}$ *				
Phenylmethanol (Benzyl alcohol)	↓	-3.82 (1.00)	-61.75 (7.78)	$<1.00x10^{-4}$ *				

(cont.) Metabolite	HG BC (J82 and Scaber) (n=30) vs LG BC (5637) (n=15)				J82 (n=15) vs Scaber (n=15)			
		ES (±ES_{SE})[a]	% variation (± % uncertainty)	*p*-value[b]		ES (±ES_{SE})[a]	% variation (± % uncertainty)	*p*-value[b]
Unknown 7					↓	-0.91 (0.73)	-12.45 (0.73)	$1.10x10^{-2}$
1-phenylethan-1-ol (1-Phenylethanol)	↑	0.67 (0.62)	12.83 (4.38)	>0.05	↑	4.10 (1.25)	41.79 (2.99)	$<1.00x10^{-4}$ *
Unknown 9								
1-phenylethan-1-one (Acetophenone)	↓	-3.94 (1.02)	-50.93 (5.77)	$<1.00x10^{-4}$ *	↓	-5.29 (1.51)	-40.58 (3.42)	$<1.00x10^{-4}$ *
Octan-1-ol (Octanol)	↑	0.78 (0.63)	14.83 (4.20)	$4.00x10^{-2}$	↑	1.55 (0.80)	26.72 (5.40)	$5.00x10^{-4}$ *
Unknown 11	↓	-3.00 (0.87)	-27.83 (2.87)	$<1.00x10^{-4}$ *				
4-Methylbenzaldehyde					↓	-1.27 (0.77)	-25.35 (8.12)	$3.00x10^{-3}$
2-phenylpropan-2-ol	↓	-2.77 (0.84)	-16.92 (2.44)	$<1.00x10^{-4}$ *	↓	-1.38 (0.78)	-6.81 (1.81)	$8.00x10^{-4}$ *
Unknown 12	↓	-0.81 (0.63)	-11.63 (4.92)	$2.00x10^{-2}$	↓	-1.11 (0.75)	-14.59 (5.01)	$7.00x10^{-3}$
Methyl benzoate					↓	-1.42 (0.78)	-25.19 (7.23)	$1.00x10^{-3}$ *
Nonanal	↓	-2.52 (0.80)	-55.34 (16.16)	$<1.00x10^{-4}$ *	↓	-1.48 (0.79)	-47.15 (14.85)	$<1.00x10^{-4}$ *
(1R,2S,5R)-5-methyl-2-(propan-2-yl)cyclohexan-1-ol (Menthol)	↓	-1.60 (0.69)	-33.93 (10.12)	$<1.00x10^{-4}$ *				
2-(1R)-4-methylcyclohex-3-en-1-ylpropan-2-ol (α-Terpineol)	↓	-2.37 (0.78)	-16.25 (2.15)	$<1.00x10^{-4}$ *				
Dodecane (N-Dodecane)	↓	-2.03 (0.74)	-17.81 (3.22)	$<1.00x10^{-4}$ *				
1,3-benzothiazole (Benzothiazole)					↓	-1.17 (0.76)	-15.20 (4.98)	$1.00x10^{-3}$ *
1-phenoxypropan-2-ol	↑	1.06 (0.65)	35.21 (6.16)	$2.00x10^{-4}$ *	↓	-2.37 (0.92)	-36.48 (6.68)	$<1.00x10^{-4}$ *
Unknown 15	↓	-1.48 (0.68)	-16.45 (3.84)	$<1.00x10^{-4}$ *				
2-hydroxy-2-methyl-1-phenylpropan-1-one	↓	-2.46 (0.79)	-13.64 (1.97)	$<1.00x10^{-4}$ *				
5-pentyloxolan-2-one (γ-Nonalactone)	↑	1.53 (0.69)	25.13 (3.53)	$<1.00x10^{-4}$ *				
Unknown 18					↑	2.54 (0.95)	50.57 (5.64)	$<1.00x10^{-4}$ *
Unknown 19	↑	1.28 (0.66)	54.10 (7.45)	$<1.00x10^{-4}$ *				
Tetradecane	↓	-1.42 (0.68)	-12.93 (7.459	$1.00x10^{-3}$ *				
(5Z)-6,10-dimethylundeca-5,9-dien-2-one (Geranylacetone)	↑	0.98 (0.64)	38.32 (7.21)	$2.00x10^{-4}$ *	↑	2.72 (0.98)	61.09 (6.11)	$<1.00x10^{-4}$ *
Hexadecane	↓	-0.66 (0.62)	-15.15 (7.18)	>0.05				
Unknown 22	↑	2.26 (0.82)	31.27 (2.69)	$<1.00x10^{-4}$ *				
Pentadecan-2-one (2-Pentadecanone)					↑	13.75 (3.55)	352.45 (3.30)	$<1.00x10^{-4}$ *
Unknown 23	↑	4.20 (1.06)	68.87 (3.46)	$<1.00x10^{-4}$ *				
Unknown 24	↑	4.04 (1.03)	60.05 (3.06)	$<1.00x10^{-4}$ *				

(cont.) Metabolite		HG BC (J82 and Scaber) (n=30) vs LG BC (5637) (n=15)			J82 (n=15) vs Scaber (n=15)		
		ES ($\pm ES_{SE}$)[a]	**% variation (± % uncertainty)**	***p*-value[b]**	**ES ($\pm ES_{SE}$)[a]**	**% variation (± % uncertainty)**	***p*-value[b]**
Unknown 25	↑	3.72 (0.98)	43.89 (2.71)	$<1.00 \times 10^{-4}$ *			

Note: BC, bladder cancer; ES, effect size; HG, high-grade cancer; LG, low-grade cancer

↑, metabolites that are increased; ↓, metabolites that are decreased in the extracellular medium

[a]ES determined as described in reference 104; values in brackets correspond to high uncertainties; [b]95% significance level (*p*-value <0.05).

*Metabolites remaining significant after Bonferroni correction, with cut off *p*-value of 1.19×10^{-3} for HG vs LG (0.05 divided by 42 metabolites) and 1.67×10^{-3} for J82 vs Scaber (0.05 divided by 30 metabolites).

5.2.2. Discriminative VOCs extracted at pH 2

Overall, and considering all comparisons between the bladder cell lines, VOCs extraction at pH 2 resulted in 32 discriminative metabolites, of which 24 were identical to those extracted at pH 7 compounds and presented the same trend of change between the cancer cell lines and the normal cell line extracellular medium (Supplementary Table S2, Appendix). Therefore, only the eight VOCs exclusively found significantly altered at pH 2 will be considered in order to avoid repetitions, except in the PCA-X and PLS-DA models, in which all relevant metabolites were included as part of the metabolic signature.

5.2.2.1. Cancer versus normal cell lines extracellular medium

For the comparison between cancer cell lines and SV-HUC-1 culture medium, seven of the eight exclusive VOCs of pH 2 extraction were found significantly altered; however, not in all cancer cell types (Table 7). For instance, there was only one VOC shared by the three cancer cell lines, namely α-methylstyrene, whose levels were diminished in the cancer cells compared with the normal cell line culture medium. As in pH 7 results, there was also metabolites shared by two of the cancer cells, according to their subtype or grade of BC. As for the subtype of BC, γ-dodecalactone and unknown 28 were increased, and unknown 27 was decreased in both TCC cell lines culture medium compared to SV-HUC-1 culture medium whereas in Scaber culture medium, these metabolites were not statistically different. In turn, benzoic acid only appeared significantly diminished in the HG cancer cells culture medium (J82 and Scaber) compared with SV-HUC-1 (Table 7). Lastly, two pH 2 VOCs were only statistically different in Scaber, namely 2-methyl-2-heptanol and methyl nonanoate, both decreased in Scaber culture medium compared with the normal one, which may represent 2 specific VOCs of the SCC cell line.

Noteworthy, and likewise 2-pentadecanone, γ-dodecalactone was not found in controls medium, being only present in all bladder cell lines culture medium, making this VOC equally important for BC metabolic signature as 2-pentadecanone.

Table 7. List of the additional VOCs, from the analysis of samples prepared at pH 2, selected as important in discriminating the cancer cell lines (5637, J82 and Scaber) from the normal cell line SV-HUC-1, which are characterized by their IUPAC (and common) name. The % variation (± % uncertainty), ES, ES_{SE} and *p*-values are represented for each VOC.

Metabolite	5637 (n=14) vs SV HUC-1 (n=15)				J82 (n=15) vs SV HUC-1 (n=15)				Scaber (n=15) vs SV HUC-1 (n=15)			
		ES (±ES_{SE})[a]	% variation (± % uncertainty)	*p*-value[b]		ES (±ES_{SE})[a]	% variation (± % uncertainty)	*p*-value[b]		ES (±ES_{SE})[a]	% variation (± % uncertainty)	*p*-value[b]
2-methyl-2-heptanol									↓	-3.95 (1.22)	-53.66 (6.60)	$<1.00x10^{-4}$ *
α-Methylstyrene	↓	-1.32 (0.77)	-17.85 (5.28)	$1.50x10^{-3}$ *	↓	-3.35 (1.10)	-33.68 (4.29)	$<1.00x10^{-4}$ *	↓	-3.30 (1.09)	-40.64 (5.50)	$<1.00x10^{-4}$ *
Benzoic acid					↓	-2.74 (0.98)	-45.41 (7.63)	$<1.00x10^{-4}$ *	↓	-7.15 (1.94)	-90.54 (8.22)	$<1.00x10^{-4}$ *
Unknown 27	↓	4.13 (1.26)	-52.07 (6.05)	$<1.00x10^{-4}$ *	↓	-5.74 (1.61)	-57.69 (5.01)	$<1.00x10^{-4}$ *				
Methyl nonanoate									↓	-1.75 (0.82)	-60.19 (17.52)	$3.00x10^{-2}$
Unknown 28	↑	4.14 (1.26)	292.71 (10.20)	$<1.00x10^{-4}$ *	↑	2.91 (1.01)	288.60 (14.42)	$<1.00x10^{-4}$ *				
5-octyloxolan-2-one (γ-dodecalactone)	↑	4.02 (1.23)	301.52 (10.63)	$<1.00x10^{-4}$ *	↑	7.04 (1.91)	309.01 (6.13)	$<1.00x10^{-4}$ *				

Note: ES, effect size

↑, metabolites that are increased; ↓, metabolites that are decreased in the extracellular medium

[a]ES determined as described in reference 104; values in brackets correspond to high uncertainties; [b]95% significance level (*p*-value <0.05).

*Metabolites remaining significant after Bonferroni correction, with cut off *p*-value of $1.52x10^{-3}$ for 5637 vs SV-HUC-1 (0.05 divided by 33 metabolites, 29 of them in common with pH 7), $1.62x10^{-3}$ for J82 vs SV-HUC-1 (0.05 divided by 31 metabolites, 26 of them in common with pH 7) and $1.39x10^{-3}$ for Scaber vs SV-HUC-1 (0.05 divided by 36 metabolites, 32 of them in common with pH 7).

In order to ascertain the robustness of the metabolic signature at pH 2 for the separation of cancer cells and the normal cell, PCA-X and PLS-DA scores scatter plots were performed considering only the set of significant metabolites. Indeed, the separation between cancer cells and the normal cell line extracellular medium was evident (figures 21a to 21f). MCCV results of the PLS-DA models (figures 21b, d and f) confirm the robustness of the metabolic signature for each cancer cell line compared with the normal one, since the classification rate is 100% (except for 5637 versus SV-HUC-1). Besides, the Q^2 values obtained for the permuted classes are inferior to the original Q^2 values (Table 8).

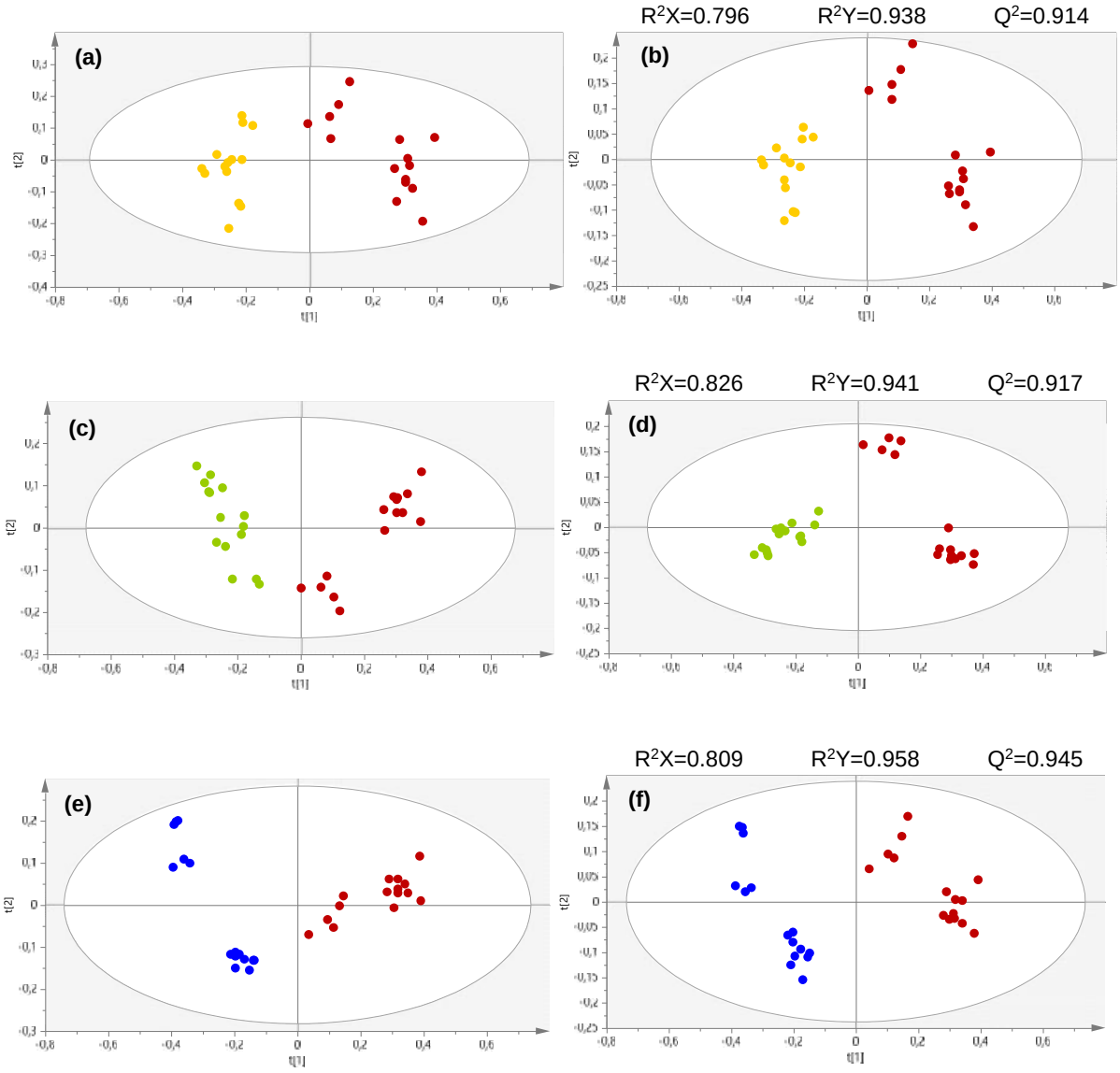

Figure 21. (a), (c) and (e) PCA-X and (b), (d) and (f) PLS-DA scores scatter plots obtained for the normal cell line (SV-HUC-1, n=15, ●) and the cancer cell lines 5637 (n=14, ●), J82 (n=15, ●) and Scaber (n=15, ●) extracellular medium, considering the discriminative metabolites only, at pH 2. The ellipses indicate the 95% confidence limit of the model.

Similarly to pH 7, boxplots of three important pH 2 metabolites responsible for the separation of the cancer cell lines and SV-HUC-1 are also represented (figure 22).

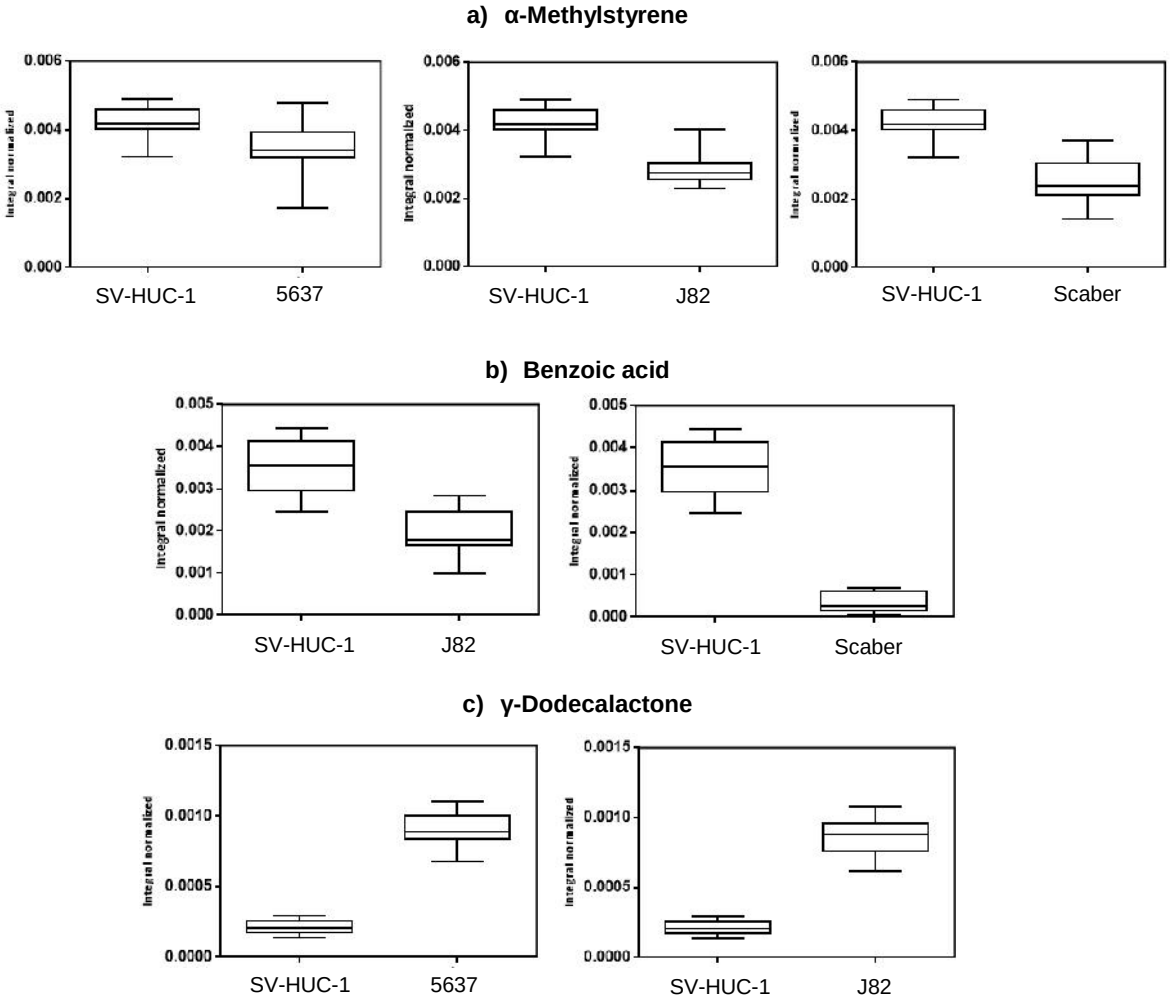

Figure 22. Boxplots for three of the VOCs found significantly altered between the normal cell line SV-HUC-1 (n=14) and the cancer cell lines 5637 (n=15), J82 (n=15) and Scaber (n=15) extracellular medium. *p*-value for a) α-methylstyrene; b) benzoic acid and c) γ-dodecalactone is $<1.00 \times 10^{-4}$. Note: *p*-values are equal in all combinations, except for α-methylstyrene, which is 1.50×10^{-3} for SV-HUC-1 *versus* 5637.

An example of a correlation network analysis of VOCs significantly altered between the extracellular medium of the cancer cell line J82 and the normal one (SV-HUC-1) was performed (figure 23). This comparison was chosen since J82 represents a higher-grade BC, thus the differences in VOCs levels between this cell line and SV-HUC-1 would be more relevant, and it represents the most common form of BC (TCC). VOCs extracted at both pHs were considered in the correlation network analysis. The explanation of the network obtained for this comparison will be further elucidated in the discussion.

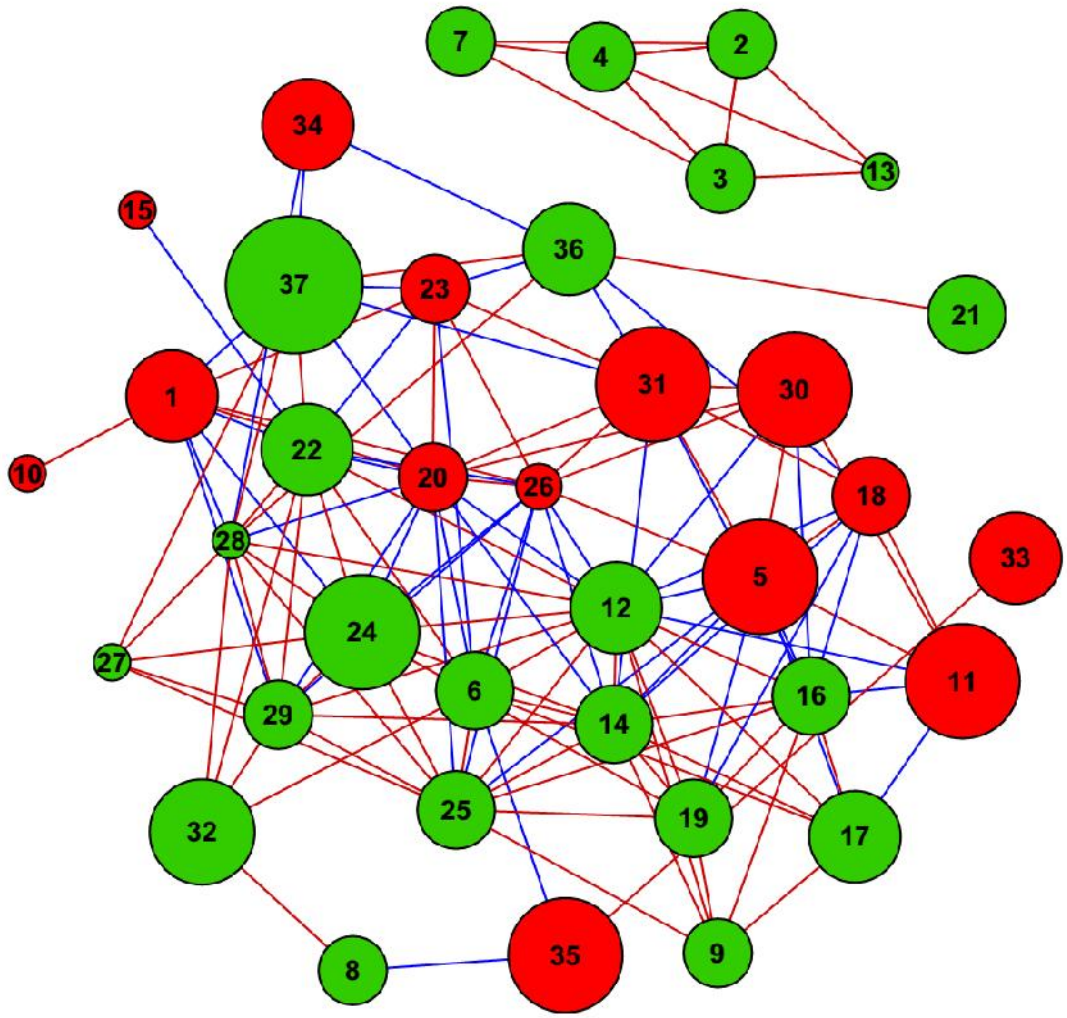

Figure 23. Correlation network of VOCs selected as important for discrimination of the cancer cell line J82 from the normal cell line SV-HUC-1 in pH 7 and 2, based on Spearman's correlation coefficients ($|r| \geq 0.8$ and $p < 0.01$). Node size denotes the effect size value of J82 in relation to SV-HUC-1. Node colours indicate direction of effect size with red for decrease and green for increase. Dark red and blue lines correspond to positive and negative correlations, respectively. For some VOCs, no correlations were found according to the threshold $|r| \geq 0.8$ and $p < 0.01$, and hence are not represented in this map, namely 2-methyl-2-butanol, cyclohexanol, unknown 3, benzyl alcohol, unknown 7, 1-octanol, nonanal, menthol, 1-phenoxypropan-2-ol, 2-undecanone and unknown 19. Compound identification: 1. Isopentanol, 2. ethylbenzene, 3. 1,3-dimethylbenzene, 4. 1,2-dimethylbenzene, 5. cyclohexanone, 6. 4-methylheptan-2-one, 7. 6-methylheptan-2-one, 8. benzaldehyde, 9. 4-methylnonane, 10. 3-ethyloctane, 11. phenol, 12. unknown 4, 13. 1,2,4-trimethylbenzene, 14. unknown 5, 15. 2-ethyl-1-hexanol, 16. unknown 8, 17. unknown 9, 18. acetophenone, 19. unknown 11, 20. 2-phenyl-2-propanol, 2. 2-nonanone, 22. unknown 14, 23. α-terpineol, 24. dodecane, 25. unknown 15, 26. 2-hydroxy-2-methyl-1-phenylpropan-1-one, 27. unknown 16, 28. unknown 17, 29. tetradecane, 30. unknown 21, 31. 2-tridecanone, 32. 2-pentadecanone, 33. α-methylstyrene, 34. benzoic acid, 35. unknown 27, 36. unknown 28, 37. γ-dodecalactone.

5.2.2.2. 5637 versus J82 and Scaber extracellular medium

Comparing 5637 (LG BC) with J82 and Scaber (HG BC), only four of the eight pH 2 exclusive VOCs were significantly altered. Levels of α-methylstyrene and γ-dodecalactone were diminished whereas (2E,4E)-deca-2,4-dienal and the unknown 28 were augmented in HG cancer cell culture medium (Table 9).

PCA-X and PLS-DA models considering the discriminative metabolites for LG BC and HG BC comparison can be observed in figure 24. As expected, separation of the cell lines was excellent. MCCV results confirm the robustness of the PLS-DA model since the set of discriminant metabolites gives a classification rate of 100% and Q^2 values obtained for the permuted classes are lower than those of the original distribution (Table 8).

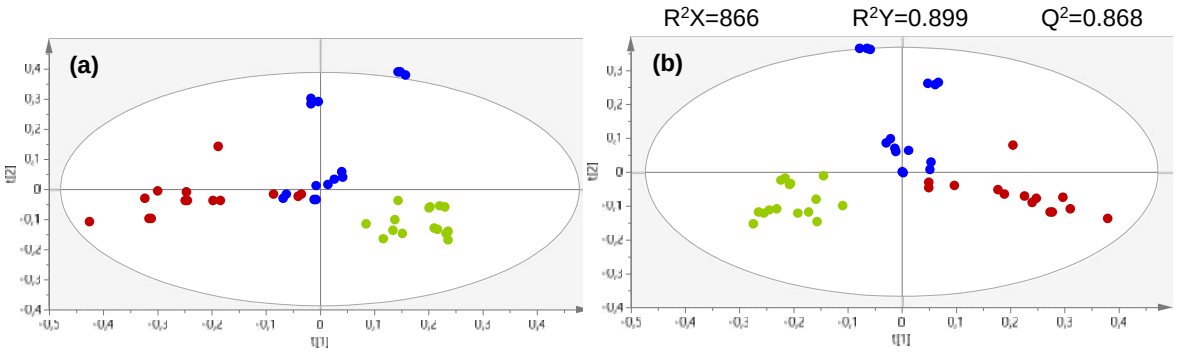

Figure 24. (a) PCA-X and (b) PLS-DA scores scatter plots obtained for the cancer cell lines 5637 (n=14, ●), J82 (n=15, ●) and Scaber (n=15, ●) extracellular medium, considering the discriminative metabolites only. The ellipses indicate the 95% confidence limit of the model.

Boxplots of two important and exclusive metabolites of pH 2 extraction responsible for the separation of the LG BC cell line and HG BC cells can be seen as well (figure 25).

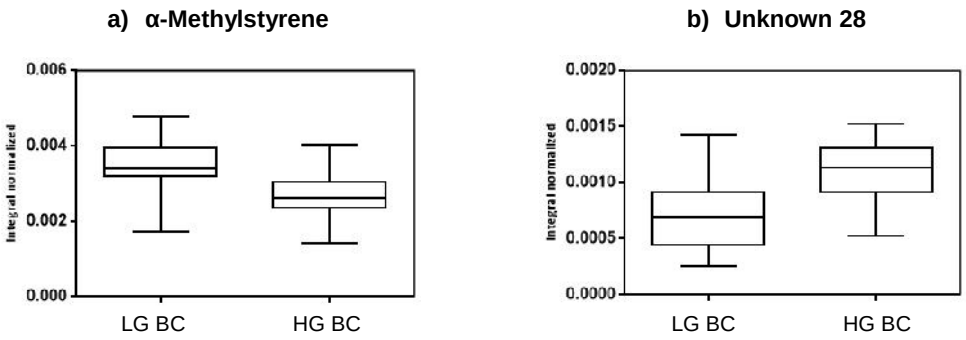

Figure 25. Boxplots for two of the VOCs found significantly altered between the LG cancer cell line 5637 (n=15) and the HG cancer cell lines J82 (n=15) and Scaber (n=15) extracellular medium. *p*-value for a) α-methylstyrene and b) unknown 28 is 5.00×10^{-4} and 6.00×10^{-4}, respectively.

5.2.2.3. J82 versus Scaber extracellular medium

The comparison of J82 and Scaber extracellular medium, at pH 2, resulted in four significantly altered VOCs, in addition to those identical with pH 7. These include benzoic acid, (2E,4E)-deca-2,4-dienal and γ-dodecalactone, which were all augmented in the extracellular medium of J82 compared to the one of Scaber (Table 9).

Additionally, PCA-X and PLS-DA models were also performed to confirm the robustness of the metabolic profile after identification of the most important metabolites (figure 26). The cancer cell lines remained clearly separated in both statistical models. MCCV results show that PLS-DA model considering the set of discriminative metabolites gives a classification rate of 100% and Q^2 values obtained for the permuted classes are lower than those of the original distribution (Table 8), confirming that the selected metabolites are responsible for the separation.

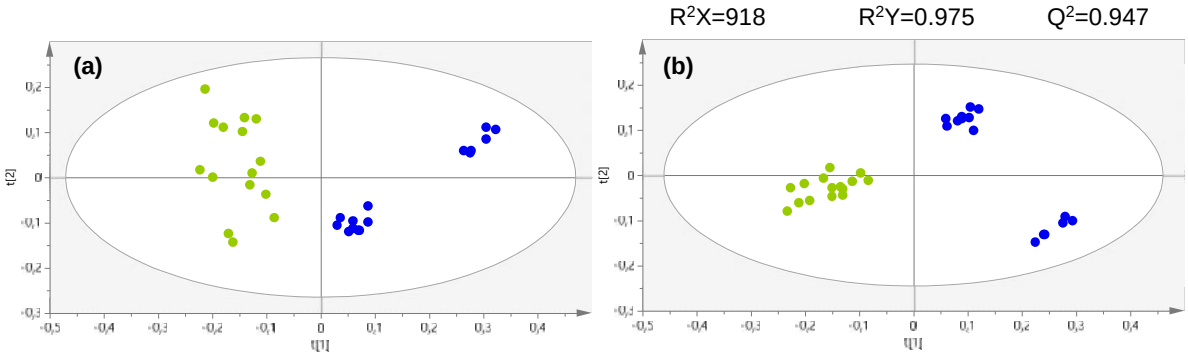

Figure 26. (a) PCA-X and (b) PLS-DA scores scatter plots obtained for the cancer cell lines J82 (n=15, ●) and Scaber (n=15, ●) extracellular medium, considering the discriminative metabolites only. The ellipses indicate the 95% confidence limit of the model.

Boxplots of three relevant pH 2 VOCs in the separation of J82 and Scaber are also represented (figure 27).

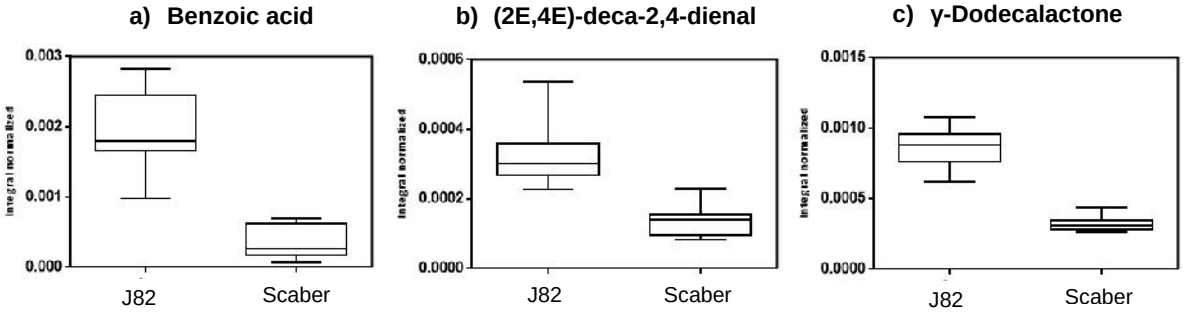

Figure 27. Boxplots for three VOCs found significantly altered between the cancer cell lines J82 (HG TCC, n=15) and Scaber (HG SCC, n=15) extracellular medium. *p*-value for a) benzoic acid; b) (2E,4E)-deca-2,4-dienal and c) γ-dodecalactone is $<1.00 \times 10^{-4}$.

Table 8. MCCV parameters of true and permuted classes obtained for pH 2 when considering the set of discriminant VOCs with statistical relevance (*p*-value < 0.05).

Models	True classes					Permuted classes				
	LV	Q^2	CR (%)	Sens. (%)	Spec. (%)	LV	Q^2	CR (%)	Sens. (%)	Spec. (%)
Set of discriminant VOCs										
5637 vs SV HUC-1	1	0.95	90	100	81	1	-0.22	49	45	52
J82 vs SV HUC-1	1	0.95	100	100	100	1	-0.23	49	49	49
Scaber vs SV HUC-1	2	0.97	100	100	100	1	-0.22	47	47	47
J82&Scaber vs 5637	2	0.94	100	100	100	1	-0.21	61	84	13
J82 vs Scaber	1	0.96	100	100	100	1	-0.21	49	49	49

Note: LV – number of latent variables, Q^2 – medium predictive power, CR – classification rate, sens. – sensitivity, spec. – specificity.

Table 9. List of the additional VOCs, from the analysis of samples prepared at pH 2, selected as important in discriminating between the cancer cell lines (5637 vs J82 vs Scaber), characterized by their IUPAC (and common) name. The % variation (± % uncertainty), ES, ES_{SE} and *p*-values are represented for each VOC.

Metabolite	HG BC (J82 and Scaber) (n=30) vs LG BC (5637) (n=14)				J82(n=15) vs Scaber (n=15)			
		ES ($\pm ES_{SE}$)[a]	% variation (± % uncertainty)	*p*-value[b]		ES ($\pm ES_{SE}$)[a]	% variation (± % uncertainty)	*p*-value[b]
α-Methylstyrene	↓	-1.41 (0.68)	-23.50 (6.66)	$5.00x10^{-4}$ *				
Benzoic acid					↑	3.61 (1.15)	476.71 (13.87)	$<1.00x10^{-4}$ *
Methyl nonanoate					↑	1.70 (0.82)	150.22 (17.92)	$2.00x10^{-2}$
(2E,4E)-deca-2,4-dienal	↑	0.62 (0.62)	33.94 (10.56)	$1.00x10^{-2}$	↑	2.96 (1.02)	131.35 (9.52)	$<1.00x10^{-4}$ *
Unknown 28	↑	1.33 (0.67)	49.75 (9.19)	$6.00x10^{-4}$ *				
5-octyloxolan-2-one (γ-dodecalactone)	↓	-0.94 (0.64)	-29.77 (11.14)	$1.00x10^{-2}$	↑	5.57 (1.57)	163.85 (5.74)	$<1.00x10^{-4}$ *

Note: BC, bladder cancer; ES, effect size; HG, high-grade cancer; LG, low-grade cancer

↑, metabolites that are increased; ↓, metabolites that are decreased in the extracellular medium

[a]E.S determined as described in reference 104; values in brackets correspond to high uncertainties; [b]95% significance level (p-value <0.05).

*Metabolites remaining significant after Bonferroni correction, with cut off *p*-value of $1.47x10^{-3}$ for HG vs LG (0.05 divided by 34 metabolites, 30 of them in common with pH 7) and $1.92x10^{-3}$ for J82 vs Scaber (0.05 divided by 26 metabolites, 22 of them in common with pH 7).

6. Discussion

Despite the potential of VOCs in discriminating cancer from non-cancer samples (98, 106, 118, 139-141), only one study has focused on BC volatilome (142) and few have reported VOCs as BC discriminative metabolites (37, 84, 87, 143). Due to the scarcity of studies investigating the BC volatilome, we decided to investigate the characteristic VOC signature of this cancer by studying the extracellular medium of three BC cell lines, which offers the advantage of not suffering interferences related to different genetic and metabolic characteristics of patients, as in the case of biological matrices, providing an unchanged metabolic signature originated from cells directly. Besides, the study of cells exometabolome was chosen since it is more suitable for VOCs collection than endometabolome, in addition to greatly reflect the cellular metabolic activity through analysis of consumed or released compounds by cells (73). Besides, VOCs released by cancer cells can be found in samples such as breath, blood or urine, which makes volatile metabolites appropriate candidates for the translatability to clinics (97).

This *in vitro* study is pivotal in investigating a VOC-biomarker panel for BC and provides a first insight on BC volatilome, which in the future may help in the early detection of this cancer, preferentially in a non-invasive way.

Two different pHs were evaluated, as well, so that it would be possible to determine the optimum pH for VOCs extraction in *in vitro* studies. The results obtained for this evaluation will be further explained. A considerable number of VOCs were significantly altered in BC cell lines compared with the normal one and between cancer cell lines themselves as well, which will also be further elucidated.

6.1. Effect of pH on the profile of VOCs released from BC culture

One of the aims of this work was to determine the optimum pH for VOCs extraction in *in vitro* approaches, so that it could be applied in future studies not only in cell metabolomics, but also in biological matrices. It is well known that alteration of the pH of the culture medium may improve the efficiency of the analysis (97), since it affects the metabolic profile of samples, as it is shown in the present work (figures 3 and 9). A neutral pH (pH 7) and an acidic one (pH 2) were evaluated in this study. Noteworthy, pH 2 was selected since acidification of urine to pH 2, being urine the best matrix to study the metabolic alterations occurring in BC (144), proved to be the optimum extraction pH in previous urinary VOCs studies as more metabolites were observed (118, 145). At first sight, pH 2 analysis resulted in more chromatographic peaks as well as more intense ones for certain compounds, such as 2-ethyl-1-hexanol and 4-methylbenzaldehyde. This

might be due to the fact that acidification of samples causes the breakdown of chemical bonds, which results in more compounds in their non-conjugated form, facilitating their detection (118, 145). Moreover, acidification of the samples predominantly favoured the formation of more acidic compounds, such as organic acids (acetic, hexanoic, heptanoic, octanoic, nonanoic and decanoic acids), benzoic acid, methyl nonanoate and methyl decanoate, as expected.

Nevertheless, after statistical analysis, it was observed that a greater number of compounds from pH 7 culture medium analysis revealed as significantly altered between cancer cell lines and the normal one, whereas, at pH 2, the number of discriminant VOCs was lower, being most of them in common with pH 7 compounds. This suggests that compounds extracted from acidic medium are not more relevant as biomarkers than those extracted from neutral medium and might explain why at pH 2, where acidic compounds are predominant over basic ones, the statistical relevant metabolites were fewer than those obtained at pH 7. An additional explanation is that, by introducing an acidic condition in cells culture medium samples, some important non-conjugated or pH-sensitive compounds may have suffered chemical degradation, which hindered their detection at pH 2.

In future works, it would be interesting to evaluate a basic pH as well, in order to compare with pH 2 and pH 7 results and to understand if basic compounds are more significant as BC biomarkers.

6.2. VOCs responsible for discriminating BC cells from SV-HUC-1

A variety of VOCs was found to be highly discriminative between cancer cells and the normal cell, in both pH 2 and pH 7. Those chemical compounds include alcohols, aldehydes, ketones, alkanes, esters, carboxylic acids and monoterpenoids. However, two major chemical classes stand out, alcohols and ketones. In the following discussion, all pH 7 metabolites and 8 more from pH 2 exclusive metabolites will be considered to avoid repetitions.

Considering the promising VOCs collected at pH 7, 17 compounds were found altered and displaying the same tendency (either increased or decreased) in all three cancer cells culture medium (Table 4), compared with the normal cell line culture medium. These results may suggest that identical metabolic pathways are altered in 5637, J82 and Scaber cell lines, such as oxidative stress or inflammation processes that lead to a common metabolic profile. From those 17 VOCs, isopentanol, cyclohexanol, 2-

phenylpropan-2-ol, 1-phenoxypropan-2-ol, 4-methylnonane, tetradecane and acetophenone have not been reported in other cancer types so far, whether in *in vitro* studies or in biological samples, which may indicate a potential specificity towards BC. As for the remaining VOCs, they were all found in previous cancer studies.

The alcohol 1-octanol, found at lower levels in BC cells culture medium, was found significantly increased in the culture medium of colorectal cancer (CRC) cells SW116 and SW480, but not in the normal cell medium (114), as opposed to the results observed in our study. It was also found diminished in urine of CRC patients, but increased in the urine of patients with leukaemia and lymphoma (98). In turn, 2-ethyl-1-hexanol was reported to be significantly released by a lung cancer cell line (NCI-H2087) (146), again contrary to what was observed for BC cell lines in this study. Furthermore, it was found elevated in melanoma tissue compared to normal tissue (99). The ketones 4-methylheptan-2-one and 2-nonanone presented significantly higher levels in the extracellular medium of BC cell lines. The first one was detected in urine of CRC patients (98). The second one was found augmented in the extracellular medium of HepG2 (liver cancer cell line) (147), as observed in this study for BC cells. Additional VOCs significantly increased in BC cells culture medium were dodecane and benzaldehyde. Similar to the results obtained for BC, dodecane was found elevated in the exhaled breath of lung cancer (39, 148) and CRC patients (149), as well as in lung cancer (39) and melanoma tissues (99) compared to healthy controls. Unlike what was observed in this study, benzaldehyde was found decreased in the extracellular medium of the liver cancer cell line HepG2 (147), which would make benzaldehyde release a characteristic of BC cells. However, it was found at higher levels in lung cancer tissue and in exhaled breath of lung cancer patients (39), and two other studies have described benzaldehyde as a potential biomarker for lung cancer diagnosis (150, 151).

Acetophenone, 1-phenoxypropan-2-ol, 1,3-dimethylbenzene and 2-pentadecanone were significantly altered in the three BC cell lines medium compared to the normal one, but in a different fashion. For instance, 1,3-dimethylbenzene and 1-phenoxypropan-2-ol were found augmented and diminished, respectively, in 5637 and J82 culture medium compared to Scaber and SV-HUC-1, which suggests metabolic differences between TCC and SCC BC. Previous studies reporting 1,3-dimethylbenzene were contradictory as this volatile was found either diminished (103) or increased (152) in exhaled breath of CRC patients. Acetophenone was found increased in 5637 but decreased in J82 and Scaber extracellular medium compared to the normal one, evidencing the influence of BC grade in its metabolism. Moreover, 2-pentadecanone was found increased in J82 but decreased in 5637, Scaber and SV-HUC-1 culture medium, making this metabolite augmentation

specific for HG TCC. In fact, this ketone was not found in controls medium (without cells), which means its origin is cellular. This makes 2-pentadecanone one of the most important and potential metabolite of BC metabolic signature. Nevertheless, two other *in vitro* studies have reported 2-pentadecanone. In one of the studies, it was observed that this VOC was highly released by lung cancer cells (A549, NCI-H446 and SK-MES-1), whereas in the normal cell (BEAS-2B) culture medium this compound was not found (122); in the other study, 2-pentadecanone was highly released by the CRC cell line SW480 (114), as well. Hence, 2-pentadecanone might be a potential biomarker but as part of a metabolite panel due to possible lack of specificity towards BC.

Additionally, some unidentified compounds were also significantly altered in cancer cells and, consequently, they are relevant in discriminating cancer from normal samples. For this reason, an attempt to identify those unknown VOCs would be important in future works with the help of standard chemicals.

The metabolite α-methylstyrene (aromatic compound) extracted at pH 2 was significantly decreased in all cancer cell lines culture medium compared to SV-HUC-1 medium (Table 7). This metabolite might also be a good candidate as BC biomarker since, to date, it has not been found altered in other diseases metabolomic studies.

In general, VOCs belonging to ketones, alkanes and aldehydes chemical classes were highly released by BC cell lines whereas alcohols were significantly decreased in BC cells extracellular medium, compared to the normal cell line medium. These results may be related with either the synthesis of fatty acids for membrane formation or β-oxidation of fatty acids for energy production, amino acids metabolism, oxidative stress and inflammation processes (147, 153, 154). For instance, one can hypothesise that 1-octanol may be associated with the synthesis of a medium-chain fatty acid, namely caprylic acid (C8:0), as 2-ethylhexanol may originate 2-ethylhexanoic acid (147); however, this requires further investigation since the source of alcohols remains uncertain. The presence of ketones in the extracellular medium of cancer cells may result from β-oxidation of branched-chain fatty acids (147, 155). The metabolism of the branched chain fatty acid nonanoic acid may result in 2-nonanone production in BC cells, as well as the metabolism of C15:0 long-chain fatty acids may originate 2-pentadecanone in J82 cells, analogous to what was demonstrated in the case of other ketone (3-heptanone) (155). It is known that cancer cells rely on β-oxidation as an alternative pathway to generate sufficient energy to sustain their higher energetic demands (156). Indeed, increase in fatty acid β-oxidation have already been associated to several cancer types, including BC (22, 87, 90, 157, 158) and others such as those of the kidney (159-161), prostate (162, 163) and pancreas (164). In an attempt to correlate the decrease in alcohols levels and the increase in

ketones, the substrates that lead to ketone formation may, in turn, be products of their respective alcohol metabolism (147), but, again, further investigation is required. Regarding simple alkanes (such as dodecane and tetradecane) and benzaldehyde, it has been described that they result from lipid peroxidation and their levels increase during inflammation and oxidative stress caused by an increase of ROS levels (98, 165-167), a hallmark of cancer. The higher oxidative stress status of cancer cells may explain the higher levels of alkanes and benzaldehyde released by BC cells in this study. Indeed, aldehydes are known biomarkers of oxidative stress and tissue damage (168, 169), which makes them good possibilities as cancer diagnosis biomarkers. The source of 1,3-dimethylbenzene or α-methylstyrene is unclear. Hence, a better understanding of the mechanisms that originate these metabolites is needed.

Some metabolites were also found significantly altered in only one of the cancer cell lines medium compared to SV-HUC-1 (Tables 4 and 7), suggesting a specificity according to the cell type characteristics and metabolic needs. Additionally, other VOCs exhibited significant differences in TCC cell lines culture medium (5637 and J82) compared to SV-HUC-1 but not in Scaber culture medium, which evidences metabolic differences between BC subtypes. Other VOCs were only statistically different in J82 and Scaber (HG BC) culture medium compared to SV-HUC-1 but not in 5637 (LG BC), emphasizing that metabolism changes as cells develop from lower grades to higher grades of cancer. Each cell line characteristic VOCs, which will be further elucidated, are also important because they may aid in the diagnosis of BC according to grade and subtype.

Regarding these results, the panel of common VOCs should be considered as a primary BC-biomarker panel for the detection of this cancer. For a more specific diagnosis, in terms of grade or even subtype, a more precise VOC panel should be employed, which will be discussed below.

6.2.1. Network analysis of the important VOCs discriminating the cell lines J82 and SV-HUC-1

A network analysis considering the VOCs identified as important in the discrimination of the cancer cell line J82 and the normal one (SV-HUC-1) was performed (figure 23) in order to get a better understanding of how VOCs would correlate and to help evaluating which are the most important VOCs to be included in BC biomarker panel. The comparison between the volatile exometabolome of J82 and SV-HUC-1 was chosen because as J82 is HG BC cell, the metabolic differences would be more significant.

Besides, J82 represents the most common subtype of BC, namely TCC, so it would be more relevant to further study the possible correlations between the altered metabolites. For these reasons, we thought it would be more interesting to carry out a network analysis of the relevant VOCs for this comparison. Of note, compounds collected at both pHs were included.

Firstly, it is quite noticeable that all benzene containing compounds are correlated with each other solely and, interestingly, 6-methylheptan-2-one correlates with ethylbenzene, 1,2-dimethylbenzene and 1,3-dimethylbenzene as well. All these VOCs are increased in J82 culture medium (Table 4) and they are positively correlated (red lines), which means that they all increase when one of the compounds increase. For instance, as ethylbenzene increases, the other 4 VOCs tend to increase as well.

Secondly, it is possible to observe a major network correlating most VOCs. For instance, as phenol decreases, acetophenone, cyclohexanone and unknown 21 also decrease as opposed to the unknowns 4, 8 and 9, which increase. Other relevant correlations involve 2-pentadecanone and γ-dodecalactone, the two VOCs that were absent in controls medium. Regarding 2-pentadecanone, it is correlated positively with benzaldehyde, 4-methylheptan-2-one, tetradecane and the unknowns 14 and 17, meaning that all increase as 2-pentadecanone increases. In turn, as γ-dodecalactone increases, the unknowns 14, 17 and 28 increase as well, in contrary to α-terpineol, benzoic acid, 2-tridecanone, 2-phenyl-2-propanol and isopentanol, whose levels diminish with the increase of γ-dodecalactone in J82 culture medium.

Moreover, VOCs with the highest differences (greater node size) are γ-dodecalactone, dodecane and 2-pentadecanone, whose levels were significantly increased in J82 culture medium, and phenol, 2-tridecanone, cyclohexanone and the unknowns 21 and 27, whose levels were significantly decreased in J82 culture medium. The node size depends on the effect size value, hence metabolites represented by the biggest nodes are associated with higher effect size values, as confirmed in tables 4 and 7. As for the others VOCs, the magnitude of change lessens as the node size diminishes. For instance, 3-ethyloctane, 2-ethyl-1-hexanol, 1,2,4-trimethylbenzene, 2-hydroxy-2-methyl-1-phenylpropan-1-one and the unknowns 16 and 17 have the smallest node sizes, suggesting that these compounds are less relevant in the discrimination of J82 from SV-HUC-1.

In conclusion, it is possible to infer that the correlations observed between the relevant volatiles, that resulted from the comparison of J82 and SV-HUC-1 culture medium, suggests a proximity on the metabolic pathways in which they are involved.

6.3. Potential discriminative VOCs of low-grade and high-grade BC

A more specific study was performed comparing the LG cell line 5637 and those of HG cancer (J82 and Scaber) in order to investigate volatiles responsible for the separation of these cell lines. Indeed, 42 metabolites obtained from pH 7 extraction and 18 from pH 2 extraction were found significantly altered, of which 14 were in common at both pHs.

Considering pH 7 results, at which more metabolites were found, the identified VOCs were mostly decreased in the extracellular medium of HG cancer cell lines (Table 6), suggesting either a higher consumption or lower production and subsequent release by both HG cancer cells compared to 5637. As for the VOCs found elevated in HG cancer cells culture medium, those with a more significant variation were benzaldehyde, 2-ethyl-1-hexanol, 1-phenoxypropan-2-ol, γ-nonalactone (cyclic ester), geranylacetone (acyclic monoterpenoid) and five unknown compounds. A possible explanation for the increase of these VOCs relies on the differences between the cells in metabolizing such compounds.

From the previous comparisons (BC cells *versus* SV-HUC-1), it was possible to identify a few VOCs that were only significantly altered in HG cancer cells culture medium compared with SV-HUC-1 culture medium, and they show the same trend of change when compared with 5637 culture medium. This suggests that the levels of those compounds change in a grade-dependent manner. Such VOCs are cyclohexanone, acetophenone, 2-hydroxy-2-methyl-1-phenylpropan-1-one, benzyl alcohol, menthol and α-terpineol, all at lower levels in HG BC (J82 and Scaber) culture medium. Apart from cyclohexanone and benzyl alcohol, none of these VOCs were reported as potential biomarkers in other cancers, suggestive of specificity towards BC. Cyclohexanone levels were found significantly higher in exhaled breath of CRC patients (149). Benzyl alcohol was also found diminished in another BC cell line (HUC T-2) culture medium compared to the normal cell line (SV-HUC-1) culture medium, but no cancer-association was made (143). In contrast to BC cells, melanoma cancer cells seem to highly release benzyl alcohol compared to normal cells (120), suggesting that cells of different cancer types present distinct metabolic needs.

The rest of the most significantly decreased volatiles in HG cancer cells extracellular medium include 6 unidentified VOCs, isopentanol, 2-methylpentan-1-ol, cyclohexanol, 2-phenylpropan-2-ol, methyl isobutyl ketone, 1,3-dimethylbenzene, styrene, 4-methylnonane, 3-ethyloctane, dodecane, tetradecane and nonanal. From these compounds, methyl isobutyl ketone, 1,3-dimethylbenzene, styrene, dodecane and nonanal were already reported as significantly changed in previous *in vitro* (39, 107, 170), *ex vivo* tissues (39) and exhaled breath cancer studies (39, 103, 149, 152, 167, 168, 171,

172). Despite lacking specificity towards BC, these relevant VOCs together with the ones possibly BC-specific, compose a potential panel of BC biomarkers that may distinguish whether it is LG or HG cancer.

Unfortunately, as far as we know, no study has investigated alterations in VOCs in earlier stages of BC compared with advanced stages that could help understand these results. Nevertheless, it is possible to conclude that HG cancer cells have different metabolic needs compared with LG cancer cells since different levels of VOCs were observed between those cells. It is known that as tumours progress and develop to more advanced stages, energetic demands of cancer cells increase in order to support DNA replication, protein synthesis and production of new components (such as those for membrane formation) so that the relentless growth and proliferation continues. Energy production is also essential for cancer cells to survive in more hostile conditions (e.g. under oxidative stress) and also to be able to migrate and adapt in other tissues (originating metastases) (30, 173). Therefore, metabolic pathways, and consequently, metabolites levels, vary in different grade cancer cells (30). Despite the lack of VOCs studies, some *in vitro* (36, 174, 175), serum (46, 88) and urinary (45, 158) studies have described metabolic alterations between LG BC and HG BC. In general, it was observed alterations in glycolysis and pyruvate metabolism (glucose and pyruvate decreased in HG cancer cells) (88, 175), in the TCA cycle (45), metabolism of amino acids (most found decreased in HG BC) (36, 46, 88) and fatty acids (particularly elevated levels of carnitines in HG BC) (45, 88, 158), and oxidative stress (GSH increased in HG BC) (174). This is in agreement with the results obtained, particularly with regard to fatty acid β-oxidation for energy production and fatty acid synthesis, important for membrane formation and cell signalling.

Hence, the results obtained in this work demonstrate that VOCs analysis can be useful for the discrimination of LG BC from advanced BC. This is important by the time of diagnosis, enabling not only the possibility of BC early diagnosis, but also the stratification of patients and monitoring of the cancer progression, which can improve the treatment options and, consequently, enhance the survival rate.

6.4. VOCs responsible for differentiating BC subtypes

Regarding pH 7 extraction, it was also possible to observe that some VOCs exhibited significant differences between the TCC cell line (J82) and SCC cell line (Scaber) in the previous comparison of BC cell lines with SV-HUC-1 extracellular medium.

This evidences that metabolic differences also vary with the subtype of BC. Indeed, those VOCs confirmed to be altered in a subtype-dependent manner after comparing J82 and Scaber culture medium, which were 1-phenoxypropan-2-ol and nonanal, both significantly diminished in J82 compared with Scaber culture medium, and 1,3-dimethylbenzene, ethylbenzene and 2-pentadecanone, highly increased in J82 culture medium (Tables 4 and 6). All these VOCs were previously found in other cancer studies (103, 109, 114, 122, 152, 167, 168, 170, 171, 176), except for 1-phenoxypropan-2-ol. Interestingly, ethylbenzene was already found elevated in the urine of BC cancer patients (142). Although no subtype or grade association was evaluated, it supports the result observed in this study regarding ethylbenzene. Again, 2-pentadecanone is of major importance due to its absence in controls medium.

Other compounds significantly altered, from pH 7 analysis, that differ between the two BC subtypes include 1,2-dimethylbenzene, styrene, 1-phenylethanol, 1-octanol, geranylacetone and unknown 18, all increased in J82 culture medium, and n-butyl acetate, cyclohexanone, 4-methylheptan-2-one, phenol, acetophenone, 2-phenylpropan-2-ol, methyl benzoate, benzothiazole and unknowns 2, 3 and 5, all decreased in J82 culture medium compared with Scaber culture medium. Among these VOCs, n-butyl acetate, cyclohexanone, 4-methylheptan-2-one and phenol were also found altered in other cancers. Cyclohexanone and 4-methylheptan-2-one have already been discussed above. As for n-butyl acetate, it was reported to be highly consumed by HepG2 (liver cancer cell line) (147) and by lung cancer cells (177, 178) due to the elevated levels of carboxylesterases (CESs) present in those cells (179), which can be responsible for converting n-butyl acetate into acetic acid and 1-butanol (147). Therefore, one can theorize that similar differences exist between J82 and Scaber; however, further investigation is required to confirm such hypothesis. To our knowledge, phenol was not reported in previous *in vitro* studies, but it has already been found at lower levels in serum (180) and urine (181) of CRC patients, and augmented in urine of breast cancer patients compared to healthy controls (110). Phenol is an intermediate of tyrosine metabolism and it might be either elevated or decreased depending on whether cells rely on tyrosine metabolism to generate energy, which might explain, in part, the differences observed for J82 and Scaber as well as in the other cancers. Altogether, this group of compounds represent possible biomarkers to distinguish TCC from SCC BC.

Despite γ-dodecalactone being a pH 2 VOC, it also demonstrated specificity towards TCC cells (Tables 7 and 9). Again, this metabolite is not present in controls medium, making this compound highly relevant for the separation of BC cells.

There are no previous studies that investigate differences in the metabolism of cells from TCC or SCC subtypes of BC, as well, in order to make more accurate interpretations of the results. Evidently, TCC and SCC cancer cell lines present distinct genetic characteristics that can be associated with up- or down-regulation of different metabolic enzymes. This leads to different metabolic reactions occurring in these two cell lines, favouring the production or consumption of certain compounds over others. Nevertheless, further investigation is required so that it is possible to understand the differences on the metabolites between the two subtypes of BC. Finding and validating biomarkers that distinguish TCC from SCC BC would be important in order to aid in the stratification of patients according to BC subtype. In fact, in BC treatment guidelines, different therapies are recommended for TCC and SCC patients (182), emphasizing the importance of an accurate diagnosis that will enable patients to benefit from the best treatment as soon as possible.

Overall, the study of bladder cells (both cancer and normal) volatilome from the extracellular medium allowed the distinction of not only cancer from normal groups, but also according to BC grade and subtype. This study, along with previous ones, showed that cancer cells can noticeably change volatile consumption or production, which is somewhat expected since cancer cells have altered metabolism. Although this study gives new insight on BC metabolism and on potential diagnosis biomarkers, further research is warranted to evaluate the robustness of the VOC profile and to validate the candidate metabolites as BC biomarkers.

Unfortunately, studies on BC-related VOCs are scarce and the origin or fate of many VOCs remains unclear, which hampers any definite conclusion to be made about the results. Therefore, it is imperative to further investigate the BC volatilome using different biological samples to better understand the metabolism of volatiles, whether they are directly tumour-derived or produced as systemic responses to inflammation, necrosis (63, 64), alteration of microbiota (11, 64) or even related to external sources (including environmental contamination (3), medication and diets (64)). This is of major importance in the identification of specific BC-related VOCs since they could improve the diagnosis and prognosis of this cancer.

Despite the difficulties and the scarcity of published works, the possibility of using VOCs as cancer biomarkers is still desired due to their relatively easy sample preparation and analysis. In BC case, the most suitable biological sample to investigate the presence of VOCs would be urine, due to its direct contact with cancer cells and to where these cells release the metabolites produced during their metabolic activity. Besides, urine has a

non-invasive collection procedure, which offers another advantage in studying this matrix in the future. Further work is needed to assess whether the VOCs found in this study may ultimately be found in urine of cancer patients in abnormal concentrations. If so, these volatiles can potentially be used as BC biomarkers, enabling the development of a non-invasive, easy-to-use diagnostic tool.

7. Conclusion and Future Perspectives

Over the last years, the interest on VOC profiles and their association with numerous diseases has increased, especially due to their non-invasive sampling approach. Indeed, VOCs have proven to be potential diagnosis biomarkers in several types of cancer such as those of the lung (176), breast (183), skin (120), colorectal (152) and prostate (103), leukaemia and lymphoma (98). Regarding *in vitro* approaches, some studies have also investigated potential VOC biomarkers for cancer diagnosis using different cancer cell types and techniques, and in all of them it was observed differences in VOCs composition between cancer and normal cell lines (39, 107, 114, 120, 121, 146, 177, 178, 183-185). Nevertheless, and to our knowledge, there are no studies on the consumption or release of VOCs by human BC cell lines, making it more difficult to compare the results obtained in this work with other data and make more specific conclusions.

One of the aims of this work was to evaluate and compare the results obtained for the extraction of VOCs at two different pHs (pH 2 and 7). Despite pH 2 analysis have resulted in more chromatographic peaks in terms of total area, pH 7 analysis resulted in a greater number of statistically significant altered metabolites. This suggests that acidic compounds are not more important as biomarkers than neutral compounds in cancer cell lines, making pH 7 the best choice for future *in vitro* metabolomic studies.

The overarching aim was, however, to discover a putative BC VOC profile. Indeed, it was observed that levels of VOCs were different when comparing the exometabolomes of BC cancer cells and the normal cell line, enabling to find:

1. A VOC biomarker panel for the early diagnosis of BC composed by 21 metabolites (isopentanol, cyclohexanol, 2-ethylhexan-1-ol, 1-octanol, 2-phenylpropan-2-ol, 1-phenoxypropan-2-ol, 4-methylnonane, dodecane, tetradecane, benzaldehyde, 4-methylheptan-2-one, acetophenone, 2-nonanone, 2-pentadecanone, 1,3-dimethylbenzene and 6 unknowns), at pH 7;
2. A VOC biomarker panel, at pH 7, for BC prognosis (prediction of disease aggressiveness) composed by 34 metabolites (isopentanol, 2-methylpentan-1-ol, cyclohexanol, 2-ethylhexan-1-ol, benzyl alcohol, 2-phenylpropan-2-ol, 1-phenoxypropan-2-ol, methyl isobutyl ketone, cyclohexanone, acetophenone, 2-hydroxy-2-methyl-1-phenylpropan-1-one, benzaldehyde, 1,3-dimethylbenzene, styrene, 4-methylnonane, 3-ethyloctane, dodecane, tetradecane, nonanal, menthol, α-terpineol, γ-nonalactone, geranylacetone and 11 unknowns);
3. A VOC biomarker panel, at pH 7, for the discrimination between two different subtypes of BC (TCC or SCC), composed by 22 volatiles (n-butyl acetate, 1,3-dimethylbenzene, ethylbenzene, 1,2-dimethylbenzene, styrene, cyclohexanone,

4-methylheptan-2-one, acetophenone, 2-pentadecanone, phenol, 1-phenylethanol, 1-octanol, 2-phenylpropan-2-ol, 1-phenoxypropan-2-ol, methyl benzoate, nonanal, benzothiazole, geranylactenone, and 4 unknowns)

4. 2-Pentadecanone and γ-dodecalactone (an exclusive metabolite of pH 2 analysis) as particularly important metabolites since they were not found in controls medium, making them potential candidates as biomarkers;
5. Ethylbenzene as a potentially good biomarker candidate, as well, since it has been reported in a previous BC urinary study (142).

These results definitely establish a proof-of-concept for the *in vitro* prediction of BC from VOCs analysis using HS-SPME/GC-MS technique and demonstrate the potential of VOCs in discriminating BC cell lines from non-cancer one. Clearly, this study cannot provide comprehensive conclusions about the results, as VOCs metabolism itself is not yet well understood. Furthermore, it should be taken into consideration that isolated *in vitro* cancer cells conditions do not reproduce those in the organism, which means that the data obtained might differ from the data obtained in biological samples. Consequently, the *in vivo* VOCs translation through analysis of human samples is mandatory so that validation of the complete BC volatilome is achieved.

In conclusion, the results are encouraging as they demonstrate the potential of VOCs as diagnosis and prognosis biomarkers, and their possible usefulness to detect BC at earlier stages. In the future, it would be worthy to explore the biological and pathological relevance of these discriminatory VOCs, as well as to attempt the identification of the unknown VOCs with the use of standards, since these compounds significantly contributed for the discrimination of BC cell lines.

In future studies, the possibility of employing 3D cultures rather than 2D ones is an exciting strategy to obtain more information since they can establish biological models more analogous to human tissues (153). Larger studies including other cancer cell types would also be interesting to evaluate the specificity of the VOC profile obtained for BC. Another interesting investigation is the search for volatiles in a simultaneous multiple matrix study, which would be advantageous to expand our knowledge about VOCs metabolism and to better compare the variations among different samples. Altogether, this might lead to the assembly of a robust, specific BC biomarker panel that would help conventional diagnosis methods. Urine would definitely be the best choice as the biological sample for the search of BC volatile biomarkers and the development of a non-invasive and relatively fast diagnostic tool, though a lot of work is still to be done.

8. REFERENCES

1. Fiehn O, Griffin J, van der Werf M, Nikolau B, Morrison NS, Goodacre R, et al. The Metabolomics Standards Initiative. Nat Biotech. 2007;25(8):846-8.

2. Gomase VS, Patil SA, Kale KV. Metabolomics. Curr Drug Metab. 2008;9(1):89-98.

3. Liu J, Litt L, Segal MR, Kelly MJ, Pelton JG, Kim M. Metabolomics of oxidative stress in recent studies of endogenous and exogenously administered intermediate metabolites. Int J Mol Sci. 2011;12(10):6469-501.

4. Ramirez T, Daneshian M, Kamp H, Bois FY, Clench MR, Coen M, et al. Metabolomics in Toxicology and Preclinical Research. Altex. 2013;30(2):209-25.

5. Fiehn O. Metabolomics-the link between genotypes and phenotypes. Plant Mol Biol. 2002;48(1-2):155-71.

6. Nagrath D, Caneba C, Karedath T, Bellance N. Metabolomics for mitochondrial and cancer studies. Biochim Biophys Acta. 2011;1807(6):650-63.

7. Rochfort S. Metabolomics Reviewed: A New "Omics" Platform Technology for Systems Biology and Implications for Natural Products Research. J Nat Prod. 2005;68(12):1813-20.

8. Lindon JC, Holmes E, Nicholson JK. So what's the deal with metabonomics? Anal Chem. 2003;75(17):384a-91a.

9. Patti GJ, Yanes O, Siuzdak G. Innovation: Metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol. 2012;13(4):263-9.

10. Griffiths WJ, Koal T, Wang Y, Kohl M, Enot DP, Deigner HP. Targeted metabolomics for biomarker discovery. Angew Chem Int Ed Engl. 2010;49(32):5426-45.

11. Davis VW, Bathe OF, Schiller DE, Slupsky CM, Sawyer MB. Metabolomics and surgical oncology: Potential role for small molecule biomarkers. J Surg Oncol. 2011;103(5):451-9.

12. Christians U, Klawitter J, Klawitter J, Brunner N, Schmitz V. Biomarkers of immunosuppressant organ toxicity after transplantation: status, concepts and misconceptions. Expert Opin Drug Metab Toxicol. 2011;7(2):175-200.

13. Niemann CU, Serkova NJ. Biochemical mechanisms of nephrotoxicity: application for metabolomics. Expert Opin Drug Metab Toxicol. 2007;3(4):527-44.

14. Serkova NJ, Reisdorph NA, Tissot van Patot MC. Metabolic markers of hypoxia: systems biology application in biomedicine. Toxicol Mech Methods. 2008;18(1):81-95.

15. Sofia M, Maniscalco M, de Laurentiis G, Paris D, Melck D, Motta A. Exploring airway diseases by NMR-based metabonomics: a review of application to exhaled breath condensate. J Biomed Biotechnol. 2011;2011:403260.

16. Oldiges M, Lütz S, Pflug S, Schroer K, Stein N, Wiendahl C. Metabolomics: current state and evolving methodologies and tools. Appl Microbiol Biotechnol. 2007;76(3):495-511.

17. German JB, Hammock BD, Watkins SM. Metabolomics: building on a century of biochemistry to guide human health. Metabolomics. 2005;1(1):3-9.

18. Zhang A, Sun H, Wang P, Han Y, Wang X. Recent and potential developments of biofluid analyses in metabolomics. J Proteomics. 2012;75(4):1079-88.

19. Rhee EP, Gerszten RE. Metabolomics and cardiovascular biomarker discovery. Clin Chem. 2012;58(1):139-47.

20. Cheng LL, Burns MA, Taylor JL, He W, Halpern EF, McDougal WS, et al. Metabolic characterization of human prostate cancer with tissue magnetic resonance spectroscopy. Cancer Res. 2005;65(8):3030-4.

21. Hakimi AA, Reznik E, Lee CH, Creighton CJ, Brannon AR, Luna A, et al. An Integrated Metabolic Atlas of Clear Cell Renal Cell Carcinoma. Cancer cell. 2016;29(1):104-16.

22. Alberice JV, Amaral AF, Armitage EG, Lorente JA, Algaba F, Carrilho E, et al. Searching for urine biomarkers of bladder cancer recurrence using a liquid chromatography-mass spectrometry and capillary electrophoresis-mass spectrometry metabolomics approach. J Chromatogr A. 2013;1318:163-70.

23. Bayet-Robert M, Loiseau D, Rio P, Demidem A, Barthomeuf C, Stepien G, et al. Quantitative two-dimensional HRMAS 1H-NMR spectroscopy-based metabolite profiling of human cancer cell lines and response to chemotherapy. Magn Reson Med. 2010;63(5):1172-83.

24. Psihogios NG, Kalaitzidis RG, Dimou S, Seferiadis KI, Siamopoulos KC, Bairaktari ET. Evaluation of tubulointerstitial lesions' severity in patients with glomerulonephritides: an NMR-based metabonomic study. J Proteome Res. 2007;6(9):3760-70.

25. Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB. Metabolomics by numbers: acquiring and understanding global metabolite data. Trends Biotechnol. 2004;22(5):245-52.

26. Spratlin JL, Serkova NJ, Eckhardt SG. Clinical applications of metabolomics in oncology: a review. Clin Cancer Res. 2009;15(2):431-40.

27. Lane AN, Fan TWM, Higashi RM. Isotopomer-Based Metabolomic Analysis by NMR and Mass Spectrometry. Methods Cell Biol. 2008;84:541-88.

28. Locasale JW, Cantley LC, Vander Heiden MG. Cancer's insatiable appetite. Nat Biotechnol. 2009;27(10):916-7.

29. Warburg O. On the origin of cancer cells. Science. 1956;123(3191):309-14.

30. Ferreira LM. Cancer metabolism: the Warburg effect today. Exp Mol Pathol. 2010;89(3):372-80.

31. Hao HX, Khalimonchuk O, Schraders M, Dephoure N, Bayley JP, Kunst H, et al. SDH5, a gene required for flavination of succinate dehydrogenase, is mutated in paraganglioma. Science. 2009;325(5944):1139-42.

32. Yun J, Rago C, Cheong I, Pagliarini R, Angenendt P, Rajagopalan H, et al. Glucose deprivation contributes to the development of KRAS pathway mutations in tumor cells. Science. 2009;325(5947):1555-9.

33. Abaffy T, Moller M, Riemer DD, Milikowski C, Defazio RA. A case report - Volatile metabolomic signature of malignant melanoma using matching skin as a control. J Cancer Sci Ther. 2011;3(6):140-4.

34. Aboud OA, Weiss RH. New opportunities from the cancer metabolome. Clin Chem. 2013;59(1):138-46.

35. Kalluri U, Naiker M, Myers MA. Cell culture metabolomics in the diagnosis of lung cancer-the influence of cell culture conditions. J Breath Res. 2014;8(2):027109.

36. Dettmer K, Vogl FC, Ritter AP, Zhu W, Nurnberger N, Kreutz M, et al. Distinct metabolic differences between various human cancer and primary cells. Electrophoresis. 2013;34(19):2836-47.

37. Pasikanti KK, Norasmara J, Cai S, Mahendran R, Esuvaranathan K, Ho PC, et al. Metabolic footprinting of tumorigenic and nontumorigenic uroepithelial cells using two-dimensional gas chromatography time-of-flight mass spectrometry. Anal Bioanal Chem. 2010;398(3):1285-93.

38. Cuperlovic-Culf M, Cormier K, Touaibia M, Reyjal J, Robichaud S, Belbraouet M, et al. H NMR Metabolomics analysis of renal cell carcinoma cells: Effect of VHL inactivation on metabolism. Int J Cancer. 2015;138(10):2439-49.

39. Filipiak W, Filipiak A, Sponring A, Schmid T, Zelger B, Ager C, et al. Comparative analyses of volatile organic compounds (VOCs) from patients, tumors and transformed cell lines for the validation of lung cancer-derived breath markers. J Breath Res. 2014;8(2):027111.

40. Beckonert O, Monnerjahn J, Bonk U, Leibfritz D. Visualizing metabolic changes in breast-cancer tissue using 1H-NMR spectroscopy and self-organizing maps. NMR Biomed. 2003;16(1):1-11.

41. Gao H, Dong B, Jia J, Zhu H, Diao C, Yan Z, et al. Application of ex vivo (1)H NMR metabonomics to the characterization and possible detection of renal cell carcinoma metastases. J Cancer Res Clin Oncol. 2012;138(5):753-61.

42. Hines KM, Ballard BR, Marshall DR, McLean JA. Structural mass spectrometry of tissue extracts to distinguish cancerous and non-cancerous breast diseases. Mol Biosyst. 2014;10(11):2827-37.

43. El-Sayed S, Bezabeh T, Odlum O, Patel R, Ahing S, MacDonald K, et al. An ex vivo study exploring the diagnostic potential of 1H magnetic resonance spectroscopy in squamous cell carcinoma of the head and neck region. Head Neck. 2002;24(8):766-72.

44. Ganti S, Weiss RH. Urine metabolomics for kidney cancer detection and biomarker discovery. Urol Oncol. 2011;29(5):551-7.

45. Wittmann BM, Stirdivant SM, Mitchell MW, Wulff JE, McDunn JE, Li Z, et al. Bladder cancer biomarker discovery using global metabolomic profiling of urine. PloS One. 2014;9(12):e115870.

46. Bansal N, Gupta A, Mitash N, Shakya PS, Mandhani A, Mahdi AA, et al. Low- and high-grade bladder cancer determination via human serum-based metabolomics approach. J Proteome Res. 2013;12(12):5839-50.

47. Chen F, Xue J, Zhou L, Wu S, Chen Z. Identification of serum biomarkers of hepatocarcinoma through liquid chromatography/mass spectrometry-based metabonomic method. Anal Bioanal Chem. 2011;401(6):1899-904.

48. Pelizzo G, Ballico M, Mimmi MC, Peiro JL, Marotta M, Federico C, et al. Metabolomic profile of amniotic fluid to evaluate lung maturity: the diaphragmatic hernia lamb model. Multidiscip Respir Med. 2014;9(1):54.

49. Wishart DS, Lewis MJ, Morrissey JA, Flegel MD, Jeroncic K, Xiong Y, et al. The human cerebrospinal fluid metabolome. J Chromatogr B. 2008;871(2):164-73.

50. Ishikawa S, Sugimoto M, Kitabatake K, Sugano A, Nakamura M, Kaneko M, et al. Identification of salivary metabolomic biomarkers for oral cancer screening. Sci Rep. 2016;6:31520.

51. Kuboniwa M, Sakanaka A, Hashino E, Bamba T, Fukusaki E, Amano A. Prediction of Periodontal Inflammation via Metabolic Profiling of Saliva. J Dent Res. 2016 Jul.

52. Sulek K, Han TL, Villas-Boas SG, Wishart DS, Soh SE, Kwek K, et al. Hair metabolomics: identification of fetal compromise provides proof of concept for biomarker discovery. Theranostics. 2014;4(9):953-9.

53. Batty CA, Cauchi M, Lourenco C, Hunter JO, Turner C. Use of the Analysis of the Volatile Faecal Metabolome in Screening for Colorectal Cancer. PloS One. 2015;10(6):e0130301.

54. Vanden Bussche J, Marzorati M, Laukens D, Vanhaecke L. Validated High Resolution Mass Spectrometry-Based Approach for Metabolomic Fingerprinting of the Human Gut Phenotype. Anal Chem. 2015;87(21):10927-34.

55. Beger RD. A review of applications of metabolomics in cancer. Metabolites. 2013;3(3):552-74.

56. Monteiro MS, Carvalho M, Bastos ML, Guedes de Pinho P. Metabolomics Analysis for Biomarker Discovery: Advances and Challenges. Curr Med Chem. 2013;20(2):257-71.

57. Weiss RH, Kim K. Metabolomics in the study of kidney diseases. Nat Rev Nephrol. 2012;8(1):22-33.

58. Chan EC, Pasikanti KK, Hong Y, Ho PC, Mahendran R, Raman Nee Mani L, et al. Metabonomic profiling of bladder cancer. J Proteome Res. 2015;14(2):587-602.

59. Saude EJ, Adamko D, Rowe BH, Marrie T, Sykes BD. Variation of metabolites in normal human urine. Metabolomics. 2007;3(4):439-51.

60. Garcia-Villalba R, Carrasco-Pancorbo A, Nevedomskaya E, Mayboroda OA, Deelder AM, Segura-Carretero A, et al. Exploratory analysis of human urine by LC-ESI-TOF MS after high intake of olive oil: understanding the metabolism of polyphenols. Anal Bioanal Chem. 2010;398(1):463-75.

61. Fernández-Peralbo MA, Luque de Castro MD. Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis. TrAC. 2012;41:75-85.

62. Gika HG, Theodoridis GA, Wilson ID. Liquid chromatography and ultra-performance liquid chromatography-mass spectrometry fingerprinting of human urine:

sample stability under different handling and storage conditions for metabonomics studies. J Chromatogr A. 2008;1189(1-2):314-22.

63. Dunn WB, Broadhurst D, Ellis DI, Brown M, Halsall A, O'Hagan S, et al. A GC-TOF-MS study of the stability of serum and urine metabolomes during the UK Biobank sample collection and preparation protocols. Int J Epidemiol. 2008;37 Suppl 1:i23-30.

64. Lauridsen M, Hansen SH, Jaroszewski JW, Cornett C. Human urine as test material in 1H NMR-based metabonomics: recommendations for sample preparation and storage. Anal Chem. 2007;79(3):1181-6.

65. Ryan D, Robards K, Prenzler PD, Kendall M. Recent and potential developments in the analysis of urine: a review. Anal Chim Acta. 2011;684(1-2):8-20.

66. Kind T, Tolstikov V, Fiehn O, Weiss RH. A comprehensive urinary metabolomic approach for identifying kidney cancer. Anal Biochem. 2007;363(2):185-95.

67. Halama A. Metabolomics in cell culture--a strategy to study crucial metabolic pathways in cancer development and the response to treatment. Arch Biochem Biophys. 2014;564:100-9.

68. Sellick CA, Hansen R, Maqsood AR, Dunn WB, Stephens GM, Goodacre R, et al. Effective quenching processes for physiologically valid metabolite profiling of suspension cultured Mammalian cells. Anal Chem. 2009;81(1):174-83.

69. Volmer M, Gettmann J, Scholz S, Buntemeyer H, Noll T. A method for metabolomic sampling of suspended animal cells using fast filtration. BMC Proc. 2011;5 Suppl 8:P93.

70. Dietmair S, Timmins NE, Gray PP, Nielsen LK, Kromer JO. Towards quantitative metabolomics of mammalian cells: development of a metabolite extraction protocol. Anal Biochem. 2010;404(2):155-64.

71. Batista U, Garvas M, Nemec M, Schara M, Veranic P, Koklic T. Effects of different detachment procedures on viability, nitroxide reduction kinetics and plasma membrane heterogeneity of V-79 cells. Cell Biol Int. 2010;34(6):663-8.

72. Dettmer K, Nurnberger N, Kaspar H, Gruber MA, Almstetter MF, Oefner PJ. Metabolite extraction from adherently growing mammalian cells for metabolomics studies: optimization of harvesting and extraction protocols. Anal Bioanal Chem. 2011;399(3):1127-39.

73. Leon Z, Garcia-Canaveras JC, Donato MT, Lahoz A. Mammalian cell metabolomics: experimental design and sample preparation. Electrophoresis. 2013;34(19):2762-75.

74. Emwas AH. The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. Methods Mol Biol. 2015;1277:161-93.

75. Dunn WB, Wilson ID, Nicholls AW, Broadhurst D. The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. Bioanal. 2012;4(18):2249-64.

76. Silva LP, Lorenzi PL, Purwaha P, Yong V, Hawke DH, Weinstein JN. Measurement of DNA concentration as a normalization strategy for metabolomic data from adherent cell lines. Anal Chem. 2013;85(20):9536-42.

77. Trygg J, Holmes E, Lundstedt T. Chemometrics in metabonomics. J Proteome Res. 2007;6(2):469-79.

78. Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics. 2013;9(2):280-99.

79. Armitage EG, Barbas C. Metabolomics in cancer biomarker discovery: current trends and future perspectives. J Pharm Biomed Anal. 2014;87:1-11.

80. Banks RE, Craven RA, Harnden P, Madaan S, Joyce A, Selby PJ. Key clinical issues in renal cancer: a challenge for proteomics. World J Urol. 2007;25(6):537-56.

81. Roessner U, Bowne J. What is metabolomics all about? Biotechniques. 2009;46(5):363-5.

82. Urquidi V, Rosser CJ, Goodison S. Molecular Diagnostic Trends in Urological Cancer: Biomarkers for Non-Invasive Diagonis. Curr Med Chem. 2012;19(22):3653-63.

83. de Gramont A, Watson S, Ellis LM, Rodon J, Tabernero J, de Gramont A, et al. Pragmatic issues in biomarker evaluation for targeted therapies in cancer. Nat Rev Clin Oncol. 2015;12(4):197-212.

84. Pasikanti KK, Esuvaranathan K, Hong Y, Ho PC, Mahendran R, Raman Nee Mani L, et al. Urinary metabotyping of bladder cancer using two-dimensional gas chromatography time-of-flight mass spectrometry. J Proteome Res. 2013;12(9):3865-73.

85. Teunissen C, Menge T, Altintas A, Alvarez-Cermeno JC, Bertolotto A, Berven FS, et al. Consensus definitions and application guidelines for control groups in cerebrospinal fluid biomarker studies in multiple sclerosis. Mult Scler. 2013;19(13):1802-9.

86. Tripathi P, Somashekar BS, Ponnusamy M, Gursky A, Dailey S, Kunju P, et al. HR-MAS NMR tissue metabolomic signatures cross-validated by mass spectrometry distinguish bladder cancer from benign disease. J Proteome Res. 2013;12(7):3519-28.

87. Putluri N, Shojaie A, Vasu VT, Vareed SK, Nalluri S, Putluri V, et al. Metabolomic profiling reveals potential markers and bioprocesses altered in bladder cancer progression. Cancer Res. 2011;71(24):7376-86.

88. Cao M, Zhao L, Chen H, Xue W, Lin D. NMR-based metabolomic analysis of human bladder cancer. Anal Sci. 2012;28(5):451-6.

89. Tate AR, Foxall PJ, Holmes E, Moka D, Spraul M, Nicholson JK, et al. Distinction between normal and renal cell carcinoma kidney cortical biopsy samples using pattern recognition of (1)H magic angle spinning (MAS) NMR spectra. NMR Biomed. 2000;13(2):64-71.

90. Huang Z, Lin L, Gao Y, Chen Y, Yan X, Xing J, et al. Bladder cancer determination via two urinary metabolites: a biomarker pattern approach. Mol Cell Prot. 2011;10(10):M111.007922.

91. Bongaerts GP, van Halteren HK, Verhagen CA, Wagener DJ. Cancer cachexia demonstrates the energetic impact of gluconeogenesis in human metabolism. Med Hypotheses. 2006;67(5):1213-22.

92. Tisdale MJ. Mechanisms of cancer cachexia. Physiol Rev. 2009;89(2):381-410.

93. Su Y-W, Chen X, Jiang Z-Z, Wang T, Wang C, Zhang Y, et al. A Panel of Serum MicroRNAs as Specific Biomarkers for Diagnosis of Compound- and Herb-Induced Liver Injury in Rats. PloS One. 2012;7(5):e37395.

94. Havrilesky LJ, Whitehead CM, Rubatt JM, Cheek RL, Groelke J, He Q, et al. Evaluation of biomarker panels for early stage ovarian cancer detection and monitoring for disease recurrence. Gynecol Oncol. 2008;110(3):374-82.

95. Laborde CM, Mourino-Alvarez L, Posada-Ayala M, Alvarez-Llamas G, Serranillos-Reus MG, Moreu J, et al. Plasma metabolomics reveals a potential panel of biomarkers for early diagnosis in acute coronary syndrome. Metabolomics. 2013;10(3):414-24.

96. Fremont RD, Koyama T, Calfee CS, Wu W, Dossett LA, Bossert FR, et al. Acute Lung Injury in Patients with Traumatic Injuries: Utility of a Panel of Biomarkers for Diagnosis and Pathogenesis. J Trauma. 2010;68(5):1121-7.

97. Schmidt K, Podmore I. Current Challenges in Volatile Organic Compounds Analysis as Potential Biomarkers of Cancer. J Biomark. 2015;2015:981458.

98. Silva CL, Passos M, Camara JS. Investigation of urinary volatile organic metabolites as potential cancer biomarkers by solid-phase microextraction in combination with gas chromatography-mass spectrometry. Br J Cancer. 2011;105(12):1894-904.

99. Abaffy T, Duncan R, Riemer DD, Tietje O, Elgart G, Milikowski C, et al. Differential volatile signatures from skin, naevi and melanoma: a novel approach to detect a pathological process. PloS One. 2010;5(11):e13813.

100. Deng C, Zhang X. A simple, rapid and sensitive method for determination of aldehydes in human blood by gas chromatography/mass spectrometry and solid-phase microextraction with on-fiber derivatization. RCM. 2004;18(15):1715-20.

101. Kwak J, Preti G. Volatile disease biomarkers in breath: a critique. Curr Pharm Biotechnol. 2011;12(7):1067-74.

102. Willis CM, Church SM, Guest CM, Cook WA, McCarthy N, Bransbury A, et al. Olfactory detection of human bladder cancer by dogs: proof of principle study. BMJ. 2004;329.

103. Peng G, Hakim M, Broza YY, Billan S, Abdah-Bortnyak R, Kuten A, et al. Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. Br J Cancer. 2010;103(4):542-51.

104. McCulloch M, Jezierski T, Broffman M, Hubbard A, Turner K, Janecki T. Diagnostic accuracy of canine scent detection in early- and late-stage lung and breast cancers. Integr Cancer Ther. 2006;5(1):30-9.

105. Pickel D, Manucy GP, Walker DB, Hall SB, Walker JC. Evidence for canine olfactory detection of melanoma. Appl Anim Behav Sci. 2004;89(1-2):107-16.

106. Di Lena M, Porcelli F, Altomare DF. Volatile organic compounds as new biomarkers for colorectal cancer: a review. Colorectal Dis. 2016;18(7):654-63.

107. Chen X, Xu F, Wang Y, Pan Y, Lu D, Wang P, et al. A study of the volatile organic compounds exhaled by lung cancer cells in vitro for breath diagnosis. Cancer. 2007;110(4):835-44.

108. Shirasu M, Touhara K. The scent of disease: volatile organic compounds of the human body related to disease and disorder. J Biochem. 2011;150(3):257-66.

109. Ulanowska A, Kowalkowski T, Trawinska E, Buszewski B. The application of statistical methods using VOCs to identify patients with lung cancer. J Breath Res. 2011;5(4):046008.

110. Silva CL, Passos M, Camara JS. Solid phase microextraction, mass spectrometry and metabolomic approaches for detection of potential urinary cancer biomarkers--a powerful strategy for breast cancer diagnosis. Talanta. 2012;89:360-8.

111. Kronert K, Kunzel M, Reutter B, Zimmermann C, Liebich HM, Luft D, et al. Urinary excretion patterns of endogenously produced alcohols in type 1 (IDDM) and type 2 (NIDDM) diabetes mellitus compared with healthy control subjects. Diabetes Res Clin Pract. 1990;10(2):161-5.

112. Nickels JD, Smith JC, Cheng X. Lateral organization, bilayer asymmetry, and inter-leaflet coupling of biological membranes. Chem Phys Lipids. 2015;192:87-99.

113. Nunes de Paiva MJ, Menezes HC, de Lourdes Cardeal Z. Sampling and analysis of metabolomes in biological fluids. Analyst. 2014;139(15):3683-94.

114. Zimmermann D, Hartmann M, Moyer MP, Nolte J, Baumbach JI. Determination of volatile products of human colon cell line metabolism by GC/MS analysis. Metabolomics. 2007;3(1):13-7.

115. Risticevic S, Niri VH, Vuckovic D, Pawliszyn J. Recent developments in solid-phase microextraction. Anal Bioanal Chem. 2009;393(3):781-95.

116. Vuckovic D, Risticevic S, Pawliszyn J. In vivo solid-phase microextraction in metabolomics: opportunities for the direct investigation of biological systems. Angew Chem Int Ed Engl. 2011;50(25):5618-28.

117. Zhang Z, Pawliszyn J. Headspace solid-phase microextraction. Anal Chem. 1993;65(14):1843-52.

118. Monteiro M, Carvalho M, Henrique R, Jeronimo C, Moreira N, de Lourdes Bastos M, et al. Analysis of volatile human urinary metabolome by solid-phase microextraction in combination with gas chromatography-mass spectrometry for biomarker discovery: application in a pilot study to discriminate patients with renal cell carcinoma. Eur J Cancer. 2014;50(11):1993-2002.

119. Vaupel P. Tumor microenvironmental physiology and its implications for radiation oncology. Semin Radiat Oncol. 2004;14(3):198-206.

120. Kwak J, Gallagher M, Ozdener MH, Wysocki CJ, Goldsmith BR, Isamah A, et al. Volatile biomarkers from human melanoma cells. J Chromatogr B. 2013;931:90-6.

121. Hanai Y, Shimono K, Oka H, Baba Y, Yamazaki K, Beauchamp GK. Analysis of volatile organic compounds released from human lung cancer cells and from the urine of tumor-bearing mice. Cancer Cell Int. 2012;12(1):7.

122. Hu YJ, Qiu YH, Chen EG, Ying KJ, Yu J, Wang P. Determination of volatile organic compounds in lung cancer cell lines and lung cancer tissue. Zhejiang Da Xue Xue Bao Yi Xue Ban. 2010;39(3):278-84.

123. Miremami J, Kyprianou N. The Promise of Novel Molecular Markers in Bladder Cancer. Int J Mol Sci. 2014;15(12):23897-908.

124. Mitra AP, Cote RJ. Molecular screening for bladder cancer: progress and potential. Nat Rev Urol. 2010;7(1):11-20.

125. Lacy S, Lopez-Beltran A, MacLennan GT, Foster SR, Montironi R, Cheng L. Molecular pathogenesis of urothelial carcinoma: the clinical utility of emerging new biomarkers and future molecular classification of bladder cancer. Anal Quant Cytol Histol. 2009;31(1):5-16.

126. Ku JH, Godoy G, Amiel GE, Lerner SP. Urine survivin as a diagnostic biomarker for bladder cancer: a systematic review. BJU int. 2012;110(5):630-6.

127. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the Folin phenol reagent. J Biol Chem. 1951;193(1):265-75.

128. Waterborg JH. The Lowry Method for Protein Quantitation. In: Walker JM, editor. The Protein Protocols Handbook. Totowa, NJ: Humana Press; 2002. p. 7-9.

129. Pragst F. Application of solid-phase microextraction in analytical toxicology. Anal Bioanal Chem. 2007;388(7):1393-414.

130. Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, et al. HMDB: the Human Metabolome Database. Nucleic Acids Res. 2007;35(Database issue):D521-6.

131. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 1999;27(1):29-34.

132. Pluskal T, Castillo S, Villar-Briones A, Oresic M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics. 2010;11:395.

133. Worley B, Powers R. Multivariate Analysis in Metabolomics. Curr Metabolomics. 2013;1(1):92-107.

134. Graca G, Duarte IF, Barros AS, Goodfellow BJ, Diaz SO, Pinto J, et al. Impact of prenatal disorders on the metabolic profile of second trimester amniotic fluid: a nuclear magnetic resonance metabonomic study. J Proteome Res. 2010;9(11):6016-24.

135. Wiklund S, Johansson E, Sjostrom L, Mellerowicz EJ, Edlund U, Shockcor JP, et al. Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. Anal Chem. 2008;80(1):115-22.

136. Westerhuis JA, van Velzen EJ, Hoefsloot HC, Smilde AK. Multivariate paired data analysis: multilevel PLSDA versus OPLSDA. Metabolomics. 2010;6(1):119-28.

137. Aickin M, Gensler H. Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods. Am J Public Health. 1996;86(5):726-8.

138. Berben L, Sereika SM, Engberg S. Effect size estimation: methods and examples. International journal of nursing studies. Int J Nurs Stud. 2012;49(8):1039-47.

139. Gasparri R, Santonico M, Valentini C, Sedda G, Borri A, Petrella F, et al. Volatile signature for the early diagnosis of lung cancer. J Breath Res. 2016;10(1):016007.

140. Mazzone PJ, Wang XF, Lim S, Choi H, Jett J, Vachani A, et al. Accuracy of volatile urine biomarkers for the detection and characterization of lung cancer. BMC Cancer. 2015;15:1001.

141. Wang Y, Hu Y, Wang D, Yu K, Wang L, Zou Y, et al. The analysis of volatile organic compounds biomarkers for lung cancer in exhaled breath, tissues and cell lines. Cancer Biomark. 2012;11(4):129-37.

142. Jobu K, Sun C, Yoshioka S, Yokota J, Onogawa M, Kawada C, et al. Metabolomics study on the biochemical profiles of odor elements in urine of human with bladder cancer. Biol Pharm Bull. 2012;35(4):639-42.

143. Pasikanti KK, Esuvaranathan K, Ho PC, Mahendran R, Kamaraj R, Wu QH, et al. Noninvasive urinary metabonomic diagnosis of human bladder cancer. J Proteome Res. 2010;9(6):2988-95.

144. Zhang J, Wei S, Liu L, Nagana Gowda GA, Bonney P, Stewart J, et al. NMR-based metabolomics study of canine bladder cancer. Biochim Biophys Acta. 2012;1822(11):1807-14.

145. Rocha SM, Caldeira M, Carrola J, Santos M, Cruz N, Duarte IF. Exploring the human urine metabolomic potentialities by comprehensive two-dimensional gas chromatography coupled to time of flight mass spectrometry. J Chromatogr A. 2012;1252:155-63.

146. Sponring A, Filipiak W, Mikoviny T, Ager C, Schubert J, Miekisch W, et al. Release of volatile organic compounds from the lung cancer cell line NCI-H2087 in vitro. Anticancer Res. 2009;29(1):419-26.

147. Mochalski P, Sponring A, King J, Unterkofler K, Troppmair J, Amann A. Release and uptake of volatile organic compounds by human hepatocellular carcinoma cells (HepG2) in vitro. Cancer Cell Int. 2013;13(1):72.

148. Handa H, Usuba A, Maddula S, Baumbach JI, Mineshita M, Miyazawa T. Exhaled breath analysis for lung cancer detection using ion mobility spectrometry. PloS One. 2014;9(12):e114555.

149. Wang C, Ke C, Wang X, Chi C, Guo L, Luo S, et al. Noninvasive detection of colorectal cancer by analysis of exhaled breath. Anal Bioanal Chem. 2014;406(19):4757-63.

150. Ligor M, Ligor T, Bajtarevic A, Ager C, Pienz M, Klieber M, et al. Determination of volatile organic compounds in exhaled breath of patients with lung cancer using solid phase microextraction and gas chromatography mass spectrometry. Clin Chem Lab Med. 2009;47(5):550-60.

151. Bajtarevic A, Ager C, Pienz M, Klieber M, Schwarz K, Ligor M, et al. Noninvasive detection of lung cancer by analysis of exhaled breath. BMC Cancer. 2009;9:348.

152. Altomare DF, Di Lena M, Porcelli F, Trizio L, Travaglio E, Tutino M, et al. Exhaled volatile organic compounds identify patients with colorectal cancer. Br J Surg. 2013;100(1):144-50.

153. Calenic B. Volatile Organic Compounds Expression in Different Cell Types: An In Vitro Approach. Int J Clin Toxicol. 2013;1:43-51.

154. Hakim M, Broza YY, Barash O, Peled N, Phillips M, Amann A, et al. Volatile organic compounds of lung cancer and possible biochemical pathways. Chem Rev. 2012;112(11):5949-66.

155. Erhart S, Amann A, Haberlandt E, Edlinger G, Schmid A, Filipiak W, et al. 3-Heptanone as a potential new marker for valproic acid therapy. J Breath Res. 2009;3(1):016004.

156. Al-Zoughbi W, Huang J, Paramasivan GS, Till H, Pichler M, Guertl-Lackner B, et al. Tumor macroenvironment and metabolism. Semin Oncol. 2014;41(2):281-95.

157. Lin L, Huang Z, Gao Y, Chen Y, Hang W, Xing J, et al. LC-MS-based serum metabolic profiling for genitourinary cancer classification and cancer type-specific biomarker discovery. Proteomics. 2012;12(14):2238-46.

158. Jin X, Yun SJ, Jeong P, Kim IY, Kim W, Park S. Diagnosis of bladder cancer and prediction of survival by urinary metabolomics. Oncotarget. 2014;5(6):1635-45.

159. Ganti S, Taylor SL, Kim K, Hoppel CL, Guo L, Yang J, et al. Urinary acylcarnitines are altered in human kidney cancer. Int J Cancer. 2012;130(12):2791-800.

160. Wettersten HI, Hakimi AA, Morin D, Bianchi C, Johnstone ME, Donohoe DR, et al. Grade-Dependent Metabolic Reprogramming in Kidney Cancer Revealed by Combined Proteomics and Metabolomics Analysis. Cancer Res. 2015;75(12):2541-52.

161. Catchpole G, Platzer A, Weikert C, Kempkensteffen C, Johannsen M, Krause H, et al. Metabolic profiling reveals key metabolic features of renal cell carcinoma. J Cell Mol Med. 2011;15(1):109-18.

162. Zha S, Ferdinandusse S, Hicks JL, Denis S, Dunn TA, Wanders RJ, et al. Peroxisomal branched chain fatty acid beta-oxidation pathway is upregulated in prostate cancer. Prostate. 2005;63(4):316-23.

163. Liu Y. Fatty acid oxidation is a dominant bioenergetic pathway in prostate cancer. Prostate Cancer Prostatic Dis. 2006;9(3):230-4.

164. Khasawneh J, Schulz MD, Walch A, Rozman J, Hrabe de Angelis M, Klingenspor M, et al. Inflammation and mitochondrial fatty acid beta-oxidation link obesity to early tumor promotion. Proc Natl Acad Sci USA. 2009;106(9):3354-9.

165. Miekisch W, Schubert JK, Noeldge-Schomburg GF. Diagnostic potential of breath analysis--focus on volatile organic compounds. Clin Chim Acta. 2004;347(1-2):25-39.

166. Phillips M, Cataneo RN, Greenberg J, Gunawardena R, Naidu A, Rahbari-Oskoui F. Effect of age on the breath methylated alkane contour, a display of apparent new markers of oxidative stress. J Lab Clin Med. 2000;136(3):243-9.

167. Poli D, Goldoni M, Corradi M, Acampa O, Carbognani P, Internullo E, et al. Determination of aldehydes in exhaled breath of patients with lung cancer by means of on-fiber-derivatisation SPME-GC/MS. J Chromatogr B. 2010;878(27):2643-51.

168. Fuchs P, Loeseken C, Schubert JK, Miekisch W. Breath gas aldehydes as biomarkers of lung cancer. Int J Cancer. 2010;126(11):2663-70.

169. Corradi M, Rubinstein I, Andreoli R, Manini P, Caglieri A, Poli D, et al. Aldehydes in exhaled breath condensate of patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2003;167(10):1380-6.

170. Bartolazzi A, Santonico M, Pennazza G, Martinelli E, Paolesse R, D'Amico A, et al. A sensor array and GC study about VOCs and cancer cells. Sensor Actuat B-Chem. 2010;146(2):483-8.

171. Poli D, Carbognani P, Corradi M, Goldoni M, Acampa O, Balbi B, et al. Exhaled volatile organic compounds in patients with non-small cell lung cancer: cross sectional and nested short-term follow-up study. Respir Res. 2005;6:71.

172. Rudnicka J, Kowalkowski T, Ligor T, Buszewski B. Determination of volatile organic compounds as biomarkers of lung cancer by SPME-GC-TOF/MS and chemometrics. J Chromatogr B. 2011;879(30):3360-6.

173. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.

174. Pendyala L, Velagapudi S, Toth K, Zdanowicz J, Glaves D, Slocum H, et al. Translational studies of glutathione in bladder cancer cell lines and human specimens. Clin Cancer Res. 1997;3(5):793-8.

175. Conde VR, Oliveira PF, Nunes AR, Rocha CS, Ramalhosa E, Pereira JA, et al. The progression from a lower to a higher invasive stage of bladder cancer is associated with severe alterations in glucose and pyruvate metabolism. Exp Cell Res. 2015;335(1):91-8.

176. Buszewski B, Ligor T, Jezierski T, Wenda-Piesik A, Walczak M, Rudnicka J. Identification of volatile lung cancer markers by gas chromatography-mass spectrometry: comparison with discrimination by canines. Anal Bioanal Chem. 2012;404(1):141-6.

177. Filipiak W, Sponring A, Filipiak A, Ager C, Schubert J, Miekisch W, et al. TD-GC-MS analysis of volatile metabolites of human lung cancer and normal cells in vitro. Cancer Epidemiol, Biomarkers Prev. 2010;19(1):182-95.

178. Sponring A, Filipiak W, Ager C, Schubert J, Miekisch W, Amann A, et al. Analysis of volatile organic compounds (VOCs) in the headspace of NCI-H1666 lung cancer cells. Cancer Biomark. 2010;7(3):153-61.

179. Imai T. Human carboxylesterase isozymes: catalytic properties and rational drug design. Drug Metab Pharmacokinet. 2006;21(3):173-85.

180. Tan B, Qiu Y, Zou X, Chen T, Xie G, Cheng Y, et al. Metabonomics identifies serum metabolite markers of colorectal cancer. J Proteome Res. 2013;12(6):3000-9.

181. Cheng Y, Xie G, Chen T, Qiu Y, Zou X, Zheng M, et al. Distinct urinary metabolic profile of human colorectal cancer. J Proteome Res. 2012;11(2):1354-63.

182. Sharma S, Ksheersagar P, Sharma P. Diagnosis and treatment of bladder cancer. Am Fam Physician. 2009;80(7):717-23.

183. He J, Sinues PM, Hollmen M, Li X, Detmar M, Zenobi R. Fingerprinting breast cancer vs. normal mammary cells by mass spectrometric analysis of volatiles. Sci Rep. 2014;4:5196.

184. Amal H, Ding L, Liu BB, Tisch U, Xu ZQ, Shi DY, et al. The scent fingerprint of hepatocarcinoma: in-vitro metastasis prediction with volatile organic compounds (VOCs). Int J Nanomedicine. 2012;7:4135-46.

185. Brunner C, Szymczak W, Hollriegl V, Mortl S, Oelmez H, Bergner A, et al. Discrimination of cancerous and non-cancerous cell lines by headspace-analysis with PTR-MS. Anal Bioanal Chem. 2010;397(6):2315-24.

9. APPENDIX

Supplementary Table S1. Results of protein quantification to assess the cellular density of all bladder cell lines (cancer and normal) in each passage and triplicates.

Cell line	Culture passage	$[protein]_{F1}$/ $mg.mL^{-1}$	$[protein]_{F2}$/ $mg.mL^{-1}$	$[protein]_{F3}$/ $mg.mL^{-1}$	$Median_{[protein]}$	Standard deviation
SV-HUC-1	6	4.31	4.63	4.36	4.43	0.14
	7	5.04	5.93	5.37	5.45	0.37
	8	4.72	4.79	5.15	4.89	0.19
	9	5.31	4.65	5.23	5.06	0.29
	10	5.86	6.26	5.24	5.79	0.42
5637	4	5.49	5.83	5.16	5.49	0.27
	5	5.06	5.08	6.02	5.39	0.45
	6	6.23	5.06	6.51	5.93	0.63
	7	5.31	5.99	5.42	5.57	0.30
	8	6.28	5.78	6.03	6.03	0.20
J82	4	5.47	4.70	4.60	4.92	0.39
	5	4.93	4.66	5.40	5.00	0.31
	6	4.98	4.84	5.73	5.18	0.39
	7	6.12	5.14	6.18	5.81	0.48
	8	4.69	5.25	5.05	5.00	0.23
Scaber	4	4.60	5.37	4.80	4.92	0.33
	5	5.10	5.94	5.55	5.53	0.34
	6	4.63	5.28	5.27	5.06	0.30
	7	5.12	5.19	4.96	5.09	0.10
	8	5.52	4.90	5.55	5.32	0.30

Note: F1, F2 and F3 represent the triplicates of each cell line passage.

Supplementary Table S2. List of VOCs selected as important in discriminating the cancer cell lines from the normal cell line SV-HUC-1 as well as in discriminating between cancer cell lines, at pH 7 and pH 2. The identification of the metabolites is based on the NIST (2014). They are characterized by their IUPAC (and common) name, characteristic ions (m/z), RT, RI_{lit}, $RI_{calc.}$, reverse match, CAs registry number, HMDB code (when available), as well as the matrices or cellular locations where they were previously found.

Metabolite	RT	Characteristic ions (m/z)	$RI_{lit.}$	$RI_{calc.}$	R. Match	CAS	HMDB code	Matrices previously found	Cellular locations
pH 7									
2-methylbutan-2-ol (2-Methyl-2-butanol) *	2.50	55+59+73	615	-	930	75-85-4	HMDB33772	NA	Cytoplasm; Extracellular
2-Pentanone	2.95	58+71+86	685	-	922	107-87-9	HMDB34235	Feces; Saliva; Urine	Cytoplasm; Extracellular
2-Methyl-2-pentanol	3.64	59	694	-	901	590-36-3	NA	NA	NA
3-methylbutan-1-ol (Isopentanol)	3.66	55+70	736	-	921	123-51-3	HMDB06007	Feces; Saliva	NA
4-methylpentan-2-one (Methyl isobutyl ketone) *	3.71	57+58+85	735	-	865	108-10-1	HMDB02939	Feces; Saliva; Urine	NA
2,3-Dimethylhexane	4.22	55+70+71+91	760	-	907	584-94-1	HMDB37617	NA	Membrane
4-Methyl-3-penten-2-one *	4.80	55+83	798	-	912	141-79-7	HMDB31563	Feces; Saliva	Cytoplasm; Extracellular
Butyl acetate (n-Butyl acetate)	5.18	56+61	812	-	844	123-86-4	HMDB31325	Feces	Cytoplasm; Extracellular
2-methylpentan-1-ol	5.67	55+56+69+70	831	-	846	105-30-6	NA	NA	NA
Ethylbenzene	6.32	91+106	855	857	938	100-41-4	HMDB59905	Blood; Feces	Membrane
1,3-dimethylbenzene (m-Xylene)	6.55	91+106	866	866	951	108-38-3	HMDB59810	Blood; Feces	Membrane
Unknown 1	6.86	57+69		878					
Cyclohexanol	7.03	57+67+82	880	885	937	108-93-0	NA	NA	NA
Ethenylbenzene (Styrene) *	7.13	77+78+82+103+104	893	889	950	100-42-5	HMDB34240	Blood; Feces; Saliva	Membrane
1,2-dimethylbenzene (o-Xylene) *	7.18	91+106	887	891	929	95-47-6	HMDB59851	Blood; Feces; Saliva	Membrane
Cyclohexanone	7.20	55+69+98	894	891	926	108-94-1	HMDB03315	Feces	NA
Unknown 2	7.85	59+95		917	-				
Unknown 3 *	8.29	59+69		929	-				
4-methylheptan-2-one *	8.43	58+59+85	943	934	922	6137-06-0	NA	NA	NA
6-methylheptan-2-one *	8.91	55+58+91	956	950	892	928-68-7	NA	NA	NA

(cont.) Metabolite	RT	Characteristic ions (m/z)	$RI_{lit.}$	$RI_{calc.}$	R. Match	CAS	HMDB code	Matrices previously found	Cellular locations
Benzaldehyde *	9.13	51+77+105+106	962	958	922	100-52-7	HMDB06115	Blood; Feces; Saliva	NA
4-Methylnonane	9.19	57+70+71	961	960	890	17301-94-9	NA	NA	NA
3-Ethyloctane	9.35	57+71	965	965	781	5881-17-4	NA	NA	NA
Phenol	9.66	65+66+94	980	976	944	108-95-2	HMDB00228	Blood; Feces; Saliva; Urine	NA
Unknown 4	9.95	58+69+85		985	-				
1,2,4-Trimethylbenzene	10.13	105+120	990	991	882	95-63-6	HMDB13733	Urine	Membrane
Unknown 5 *	10.33	57+59+103		998	-				
Unknown 6 *	10.62	57+59+69+71+89		1008	-				
2,6-Dimethylnonane *	10.70	57+70+71	1018	1011	927	17302-28-2	NA	NA	NA
2-ethylhexan-1-ol (2-Ethyl-1-hexanol) *	11.22	55+56+57+70+83	1030	1027	922	104-76-7	HMDB31231	Feces; Saliva	Extracellular; Membrane
Phenylmethanol (Benzyl alcohol)	11.34	55+77+79+107+108	1036	1031	926	100-51-6	HMDB03119	Feces; Saliva	NA
Unknown 7	11.72	57+69+73		1044	-				
Unknown 8	11.90	58+71+84		1050	-				
1-phenylethan-1-ol (1-Phenylethanol)	12.13	51+77+79+107	1061	1057	926	98-85-1	HMDB32619	Feces; Saliva	Cytoplasm; Extracellular
Unknown 9	12.19	55+69+97		1059	-				
1-phenylethan-1-one (Acetophenone) *	12.26	51+77+105	1065	1062	939	98-86-2	HMDB33910	Feces; Saliva	Cytoplasm; Extracellular
Octan-1-ol (Octanol)	12.45	55+56+69+70	1071	1068	852	111-87-5	HMDB01183	Feces; Urine	Extracellular; Membrane
Unknown 10	12.51	55+59		1070	-				
Unknown 11	12.65	55+57+69+70+71+83		1075	-				
4-Methylbenzaldehyde *	12.80	65+91+119+120	1079	1080	933	104-87-0	HMDB29638	Feces; Saliva	Cytoplasm; Extracellular
2-phenylpropan-2-ol	12.89	121	1090	1082	891	617-94-7	NA	NA	NA
Unknown 12	12.98	67+123+138	1092	1085	-	504-20-1	NA	NA	NA
2-Nonanone *	13.05	57+58+59+71	1092	1088	897	821-55-6	HMDB31266	Feces; Saliva; Urine	Membrane
Methyl benzoate *	13.15	51+77+105	1094	1091	904	93-58-3	HMDB33968	NA	Cytoplasm; Extracellular
Nonanal	13.47	55+56+57+70+81	1104	1102	898	124-19-6	HMDB59835	Blood; Feces; Saliva	Membrane

(cont.) Metabolite	RT	Characteristic ions (m/z)	$RI_{lit.}$	$RI_{calc.}$	R. Match	CAS	HMDB code	Matrices previously found	Cellular locations
Unknown 13	14.61	55+56+69+97		1140	-				
Unknown 14	15.48	58+69+70+85		1170	-				
(1R,2S,5R)-5-methyl-2-(propan-2-yl)cyclohexan-1-ol (Menthol) *	15.64	55+71+81+95	1175	1175	922	2216-51-5	HMDB03352	Blood; Feces; Saliva; Urine	Extracellular; Membrane
Naphthalene	15.84	91+119+128	1182	1182	900	91-20-3	HMDB29751	Feces; Saliva	Membrane
2-[(1R)-4-methylcyclohex-3-en-1-yl]propan-2-ol (α-Terpineol) *	16.14	55+59+81+93+121	1189	1192	883	98-55-5	HMDB04043	Feces	Extracellular; Membrane
Dodecane (N-Dodecane)	16.32	57+71+85	1200	1198	888	112-40-3	HMDB31444	Feces; Saliva	Membrane
1,3-benzothiazole (Benzothiazole)	16.97	69+108+135	1229	1222	921	95-16-9	HMDB32930	NA	Cytoplasm; Extracellular
1-phenoxypropan-2-ol	17.50	77+94	1246	1241	929	770-35-4	NA	NA	NA
Unknown 15 *	17.66	57+91+175		1246	-				
2-hydroxy-2-methyl-1-phenylpropan-1-one *	18.52	57+59+77	1278	1277	746	7473-98-5	NA	NA	NA
Undecan-2-one (2-Undecanone)	18.86	58+71	1294	1289	782	112-12-9	HMDB33713	Feces; Saliva	Membrane
Unknown 16	19.36	55+57+69+83	(97)	1307	-				
Unknown 17	19.58	55+57+69+85		1316	-				
5-pentyloxolan-2-one (γ-Nonalactone)	20.62	85	1363	1355	854	104-61-0	NA	NA	NA
Unknown 18	21.46	55+69+83+91+119		1387	-				
Unknown 19	21.64	55+69+85		1394	-				
Tetradecane	21.76	57+71+85	1400	1398	915	629-59-4	HMDB59907	Feces; Saliva	Membrane
Dodecanal (Lauric aldehyde)	21.96	55+57+67+69+82	1409	1406	930	112-54-9	HMDB33933	Feces; Saliva	Extracellular; Membrane
(5Z)-6,10-dimethylundeca-5,9-dien-2-one (Geranylacetone)	22.91	69+91+119	1453	1444	859	3796-70-1	HMDB31846	NA	Extracellular; Membrane
Dodecan-1-ol (1-Dodecanol)	23.60	55+56+57+69+70+83	1473	1471	894	112-53-8	HMDB11626	Feces; Saliva	Extracellular; Membrane
Unknown 20	23.87	55+91+105+115+129		1482	-				
Unknown 21	24.01	91+105		1488	-				
Tridecan-2-one (2-Tridecanone) *	24.10	58+59+91+119	1497	1491	779	593-08-8	HMDB34148	Feces; Saliva	Membrane
Hexadecane *	26.64	57+71+85	1600	1598	908	544-76-3	HMDB33792	Feces; Saliva	Membrane
Unknown 22	28.63	77+81+99+107+121		1683	-				
Pentadecan-2-one (2-Pentadecanone) *	28.80	58+59+71	1698	1690	854	2345-28-0	HMDB31081	Saliva	Membrane

(cont.) Metabolite	RT	Characteristic ions (m/z)	$RI_{lit.}$	$RI_{calc.}$	R. Match	CAS	HMDB code	Matrices previously found	Cellular locations
Unknown 23	29.45	107+135		1718	-				
Unknown 24	29.52	91+107+121+163		1724	-				
Unknown 25	29.64	107+121+149		1730	-				
Unknown 26	29.87	104+107+135		1741	-				
pH 2									
2-methylheptan-2-ol	7.26	55+57+59	885	894	783	625-25-2	NA	NA	NA
Prop-1-en-2-ylbenzene (α-Methylstyrene)	9.74	59+78+103+117+118	976	978	835	98-83-9	HMDB59899	NA	Membrane
Benzoic acid	15.41	51+77+105+122	1170	1168	927	65-85-0	HMDB01870	Blood; Feces; Saliva; Urine	Cytoplasm; Extracellular; ER
Unknown 27	16.57	55+59+61+89		1207	-				
Methyl nonanoate	17.09	55+73+74+87+147	1225	1226	722	1731-84-6	HMDB31264	NA	Extracellular; Membrane
(2E,4E)-deca-2,4-dienal	19.55	81	1317	1315	885	2363-88-4	NA	NA	NA
Unknown 28	22.17	55+57+71+73+74+87		1414	-				
5-octyloxolan-2-one (γ-dodecalactone)	28.31	85	1678	1669	882	2305-05-7	NA	NA	NA

Note: CAS, Chemical Abstracts Service; HMDB, Human Metabolome Database, NA, not available; R. Match, reverse match from NIST (2014); $RI_{lit.}$ average retention index according to literature; $RI_{calc.}$ retention index calculated based on the *n*-alkanes series; RT, retention time

*Metabolites that are also altered at pH 2.